高等教育“十三五”规划教材

姜本源 屠良平 张金海 宋介珠 主编

概率论与数理统计

（第2版）

清华大学出版社

高等教育"十三五"规划教材

概率论与数理统计

（第 2 版）

姜本源　屠良平　张金海　宋介珠　主编

清 华 大 学 出 版 社
北　京

内 容 简 介

本书共 8 章，分为两部分：概率论部分(第 1～5 章)主要讲述了随机事件、一维及多维随机变量的分布、随机变量的数字特征、大数定律及中心极限定理等内容；统计部分(第 6～8 章)主要讲述了区间估计与假设检验两种统计推断方法，并简单介绍了方差分析与回归分析.

本书适用于非数学类的本专科生，也可供相关工程技术人员使用.

图书在版编目(CIP)数据

概率论与数理统计/姜本源等主编. —2 版. —北京：清华大学出版社，2018
ISBN 978-7-302-50260-9

Ⅰ. ①概…　Ⅱ. ①姜…　Ⅲ. ①概率论 ②数理统计　Ⅳ. ①O21

中国版本图书馆 CIP 数据核字(2018)第 114716 号

责任编辑：佟丽霞
封面设计：常雪影
责任校对：王淑云
责任印制：丛怀宇

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市铭诚印务有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：11.5　　**字　　数**：280 千字
版　　次：2010 年 3 月第 1 版　2018 年 9 月第 2 版　　**印　　次**：2018 年 9 月第1 次印刷
定　　价：29.00 元

产品编号：068399-01



本书编者依据教育部制定的"工科类本科数学基础课程教学基本要求",针对使用本教材的同行和广大读者的反馈信息,总结第1版教材在教学使用中的经验和不足,运用现代概率统计学的观点、方法,对其进行重新设计、修订.

本书保留了第1版的体系,在突出知识结构严谨,简明易学特色的前提下,对第1版内容作了如下修改:第1章"概率论的基本概念"改名为"随机事件及其概率";为便于读者学习,在每章都增加了"小结",梳理本章的主要内容和知识结构,指明重点、难点;对内容、例题、习题作了适当的增减;删去了原附录中"统计软件Minitab简介"的内容,在附录中增加了"常用术语的汉英对照",以便读者查阅、学习相关英文文献资料.

本书第1章、第8章由姜本源负责修改,第2章、第3章由宋介珠负责修改,第4章、第5章由张金海负责修改,第6章、第7章由屠良平负责修改,马东怡、丁桂艳、卢飞龙负责习题部分的修订,马东怡负责附录B的编写,全书由姜本源统稿.本次修订得到了清华大学出版社和辽宁科技大学有关领导的大力支持,同时也得到了广大师生的关心和帮助,在此深表感谢.

由于编者水平有限,疏漏和不妥之处在所难免,恳请专家和广大师生提出宝贵意见,我们将与时俱进,不断改进提高.

编者

2018年1月



本书是作者在积累多年的工科概率论与数理统计课程教学经验并参考大量的国内外同类教材基础上编写的.在教材体系与内容安排上,我们注重联系工科各专业的实际,将收集的工程与生活中概率统计的应用实例写入教材,旨在培养学生把数学理论应用于工程实践的意识与能力,以适用于应用型人才的培养.本书在确保知识结构严密的基础上,力图概念讲述简洁明了,理论论证深入浅出.

全书共8章,分为两部分:概率论部分主要讲述了随机事件、一维及多维随机变量的分布、随机变量的数字特征、大数定律及中心极限定理等内容;统计部分主要讲述了区间估计与假设检验两种统计推断方法,并简单介绍了方差分析与回归分析.最后在本书的附录中扼要介绍了 Minitab 软件的使用方法.

本书在习题配备上,力求突出工科特色.每章的习题分成两部分:一部分是基本要求题;另一部分是补充与提高题.

本书适用于工科本、专科生的概率论与数理统计课程教学(其中“*”部分是供选学的).

本书的主编为王金萍、张金海、姜本源、宋介珠,副主编为张大庆、刘昊、孟丽新、丁桂艳,全书由何希勤教授统稿.在本书的编写过程中得到了辽宁科技大学理学院领导及同行们的大力支持,在此表示感谢.

由于编者水平有限,书中难免有不妥之处,恳请专家及广大师生提出宝贵意见.

编者

2010年1月




第 1 章

随机事件及其概率

对于自然界和日常生活中发生的各种现象,通常可以将其分为两类.一类称为确定性现象,即在一定条件下必然发生或必然不发生的现象.例如:太阳从东方升起;上抛的石子一定落下.另一类称为随机现象,即在一定条件下可能发生,也可能不发生的现象,其结果无法事先预知.例如:掷一枚硬币落在平面上,可能字面朝上,也可能另一面朝上;远距离射击一个目标,可能击中,也可能击不中.尽管随机现象具有不确定性,但人们通过进一步研究却发现这些随机现象在大量重复试验下,其结果具有某种规律性.例如,多次重复抛一枚硬币得到字面朝上的次数大体上占一半.这种在大量重复试验中,随机现象所呈现出的内在规律性称为统计规律性.因而,随机现象具有两个特点:①在一次具体试验中,现象可能发生,也可能不发生,即结果呈现出不确定性;②在大量重复试验中,其结果具有统计规律性.

概率论与数理统计就是研究随机现象的统计规律的一门数学学科.作为数学的一个重要分支,概率论与数理统计的理论和方法在自然科学、工程技术和社会科学等领域有着广泛的应用.概率论与数理统计与其他学科的相互渗透融合是现代科学技术发展的特征之一.

1.1 随机事件

1.1.1 随机试验

若试验具有以下特点:

(1) 在相同条件下可以重复进行;

(2) 试验的结果不止一个,但事先能明确其所有的可能结果;

(3) 每次试验之前,不能确定会出现哪个结果.

则称该试验为**随机试验**,简称**试验**.随机试验通常用字母 E 表示.本书中后面提到的试验都是指随机试验,我们就是通过研究随机试验来研究随机现象的.

例 1 下列试验都是随机试验.

E_1:掷一只骰子,观察朝上的那一面的点数.

E_2:在一批产品中,任取一件产品,观察其是正品还是次品.

E_3:射击同一个目标,直到击中为止,记录射击次数.

E_4:记录电话交换台在一分钟内接到的呼叫次数.

E_5:从一批灯泡中,任取一只,测量其寿命.

1.1.2 随机事件

1. 样本空间

尽管在每次具体试验之前不能预知试验的结果，但试验的所有可能结果是已知的. 我们将试验 E 的所有可能结果组成的集合称为试验 E 的**样本空间**，记为 S. 样本空间中的元素，即试验 E 的每个结果，称为**样本点**.

例 2 写出例 1 中各随机试验的样本空间.

$S_1=\{1,2,3,4,5,6\}$；

$S_2=\{$正品，次品$\}$；

$S_3=\{1,2,3,4,\cdots\}$；

$S_4=\{0,1,2,3,\cdots\}$；

$S_5=\{t\,|\,t\geqslant 0\}$.

2. 随机事件

试验 E 的样本空间 S 的子集称为试验 E 的**随机事件**，简称**事件**. 随机事件用大写英文字母 A,B,C 等表示. 在一次试验中，若该子集中的一个样本点出现，则称这一事件**发生**. 在一次具体试验中，一个事件可能发生也可能不发生.

特别地，只含一个样本点的集合，称为**基本事件**. 例如例 1 中的试验 E_1 有 6 个基本事件 $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$；试验 E_3 有无数个基本事件 $\{1\},\{2\},\{3\},\cdots$.

样本空间 S 是其自身的子集，也是一个事件，但其在每次试验中都必然发生，称为**必然事件**. 空集 $\varnothing$ 也是样本空间的子集，其在每次试验中都不发生，称为**不可能事件**. 必然事件和不可能事件的发生与否，已经失去了“不确定性”，因而本质上它们已经不属于随机事件；但是为了讨论问题方便起见，我们还是把它们作为随机事件来处理，可以将其理解为随机事件的两个极端情况.

1.1.3 事件之间的关系和运算

概率论的任务之一，就是研究随机事件的统计规律；通过对较为简单的事件的规律的研究去掌握更为复杂的事件的规律，因此需要研究事件之间的相互关系和运算. 随机事件是一个集合，所以事件之间的关系和运算与集合之间的关系和运算是完全类似的. 20 世纪 30 年代初，冯・米泽斯(Von Mises)率先用集合论的观点来研究随机事件，为概率论的公理化研究作出了贡献.

1. 包含关系

若事件 A 发生必然导致事件 B 发生，则称事件 B **包含**事件 A(或称事件 A **包含于**事件 B)，记为 $B\supset A$(或 $A\subset B$). 任何事件都包含于样本空间 S.

2. 相等关系

若事件 B 包含事件 A，且事件 A 又包含事件 B，即 $B\supset A$ 且 $A\supset B$，则称事件 A 与事件 B **相等**，记为 $A=B$. 此时事件 A 与事件 B 所包含的样本点相同.

3. 事件的并(和)

若事件 A 和事件 B 至少有一个发生，这一事件称为事件 A 与事件 B 的**并**(或**和**)，记为

$A\cup B$. 事件 $A\cup B$ 也称为事件 A 与事件 B 的**和事件**. 称 $\bigcup\limits_{k=1}^{n}A_k$ 为 n 个事件 $A_1,A_2,\cdots,A_n$ 的和事件;称 $\bigcup\limits_{k=1}^{\infty}A_k$ 为可列个事件 $A_1,A_2,A_3,\cdots$的和事件.

4. 事件的交(积)

若事件 A 和事件 B 同时发生,这一事件称为事件 A 与事件 B 的**交**(或**积**),记为 $A\cap B$(或 AB). 事件 $A\cap B$ 也称为事件 A 与事件 B 的**积事件**. 称 $\bigcap\limits_{k=1}^{n}A_k$ 为 n 个事件 $A_1,A_2,\cdots,A_n$ 的积事件;称 $\bigcap\limits_{k=1}^{\infty}A_k$ 为可列个事件 $A_1,A_2,A_3,\cdots$的积事件.

5. 互不相容

若事件 A 与事件 B 不能同时发生,即 $A\cap B=\varnothing$,则称事件 A 与事件 B 是**互不相容的**(或**互斥的**). 任何两个不同的基本事件是互不相容的.

6. 逆事件

若事件 A 与事件 B 互不相容,且在任何一次试验中二者必有一个发生,即 $A\cap B=\varnothing$ 且 $A\cup B=S$,则称事件 A 与事件 B 是**互逆的**(或**相互对立的**). 亦称事件 A 与事件 B 互为**逆事件**(或**对立事件**). 事件 A 的逆事件记为 $\overline{A}$.

7. 事件的差

若事件 A 发生而事件 B 不发生,这一事件称为事件 A 与事件 B 的**差**,记为 $A-B$. 此时也可以看成事件 A 与事件 $\overline{B}$ 的交,即 $A-B=A\overline{B}$.

英国逻辑学家文(Venn)提供了一种用直观的几何解释来表示事件之间的各种关系的几何图形,称为**文氏图**(见图 1-1).

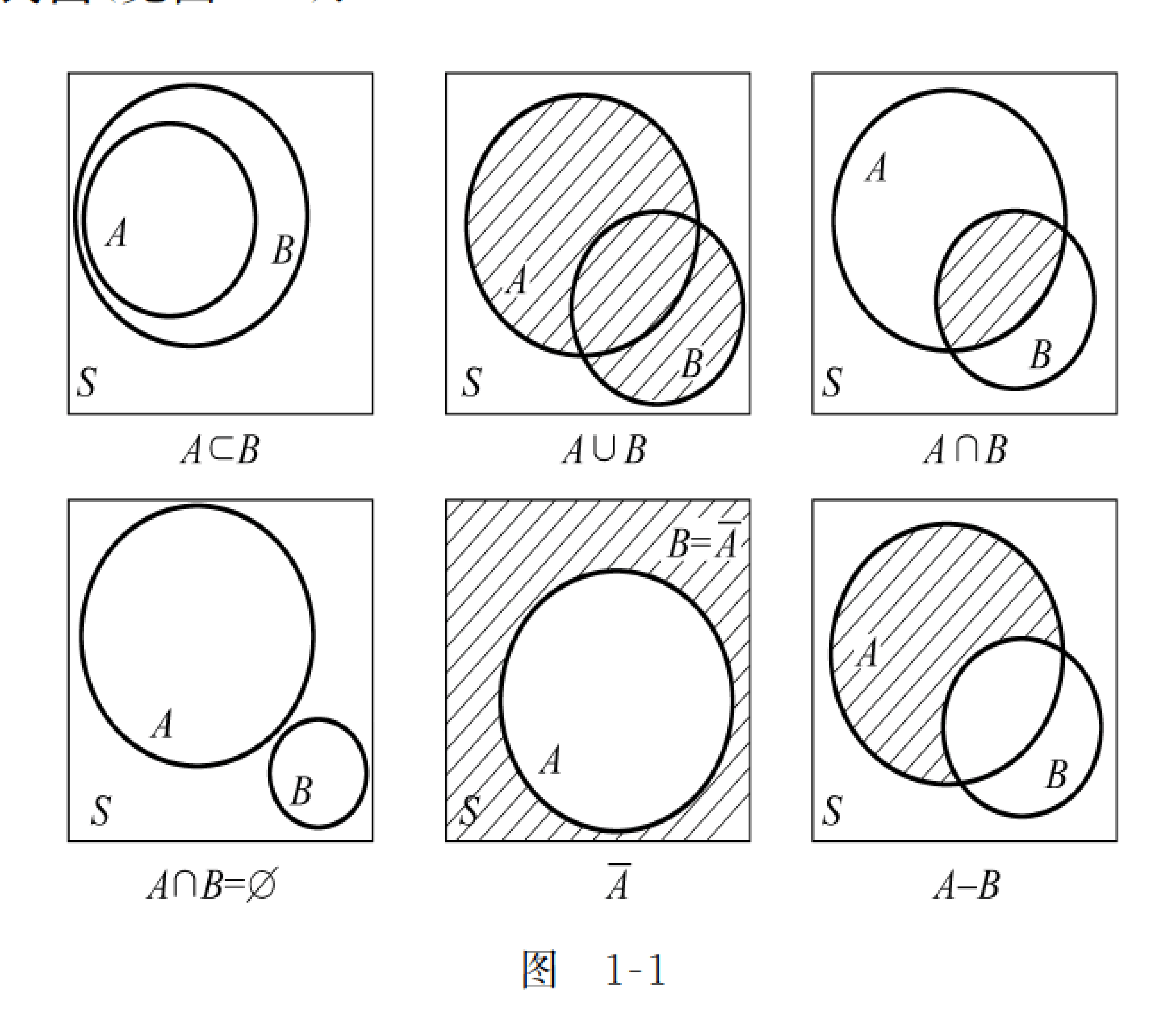


图 1-1

为了便于理解,我们把事件的关系与运算的两种解释的对照列表如下(见表 1-1).

事件运算具有如下基本性质:

交换律: $A\cup B=B\cup A$, $A\cap B=B\cap A$.

表　1-1

符　号	概率论的解释	集合论的解释
$A\subset B$	事件 A 发生必然导致事件 B 发生	集合 A 是集合 B 的子集
$A=B$	事件 A 与事件 B 相等	集合 A 与集合 B 相等
$A\cup B$	事件 A 和事件 B 至少有一个发生	集合 A 与集合 B 的并集
$A\cap B$	事件 A 和事件 B 同时发生	集合 A 与集合 B 的交集
$AB=\varnothing$	事件 A 与事件 B 不能同时发生	集合 A 与集合 B 无公共元素
$\overline{A}$	事件 A 的逆事件(对立事件)	集合 A 的补集
$A-B$	事件 A 发生而事件 B 不发生	集合 A 与集合 B 的差集

结合律：$(A\cup B)\cup C=A\cup(B\cup C)$，$(A\cap B)\cap C=A\cap(B\cap C)$.

分配律：$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$，

$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$.

德·摩根(De Morgan)律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

例 3　设 A,B,C 为三个事件，一些事件可表示如下.

(1) 三个事件都发生：ABC.

(2) 三个事件中至少发生一个：$A\cup B\cup C$.

(3) A 发生而 B 与 C 都不发生：$A\overline{B}\overline{C}$ 或 $A-B-C$ 或 $A-(B\cup C)$.

(4) A 与 B 发生而 C 不发生：$AB\overline{C}$ 或 $AB-C$.

(5) 三个事件恰好发生一个：$A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C$.

(6) 三个事件恰好发生两个：$AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$.

(7) 三个事件全不发生：$\overline{A}\overline{B}\overline{C}$ 或 $\overline{A\cup B\cup C}$.

(8) 三个事件中至少有一个不发生：$\overline{A}\cup\overline{B}\cup\overline{C}$ 或 $\overline{ABC}$.

1.1.4　排列与组合

为了便于学习概率的计算，现简要介绍排列与组合的相关知识.

从一集合中有次序地选取一组元素称为**排列**. 从一集合中选取一组元素而不计其次序称为**组合**. 排列组合的计算基于如下两个原理.

1. 加法原理

若事物 A 有 m 种取法，事物 B 有另外 n 种取法，则事物 A 与事物 B 任取其一共有 $m+n$ 种取法. 这里，事物 A 与事物 B 的选取是“平行”的，具有“互斥性”.

2. 乘法原理

若事物 A 有 m 种取法，而事物 B 又有另外 n 种取法，则先事物 A 后事物 B 共有 mn 种取法. 这里，事物 A 与事物 B 的选取是有“次序”的，选取过程中存在“依赖关系”.

常用计算结果如下：

从 n 个不同的元素中取 r 个不同元素的排列数为 $\mathrm{P}_n^r=\mathrm{A}_n^r=n(n-1)(n-2)\cdots(n-r+1)$.

从 n 个不同的元素中取 r 个元素作排列，但选取过程中任何元素均允许重复出现，此时排列数为 n^r.

从 n 个不同的元素中取 r 个不同元素的组合数为

$$\binom{n}{r}=C_n^r=\frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}=\frac{P_n^r}{r!}=\frac{n!}{r!(n-r)!}.$$

从 n 个不同的元素中取 r 个元素作组合,但选取过程中任何元素均允许重复出现,此时组合数为

$$\binom{n+r-1}{r}.$$

从 n 个不同的元素中取 k 个部分,第一部分 r_1 个元素,第二部分 r_2 个元素,…,第 k 部分 r_k 个元素,且满足 $r_1+r_2+\cdots+r_k=n$,此时不同的取法数为

$$\frac{n!}{r_1!r_2!\cdots r_k!}.$$

1.2 随机事件的概率

在随机试验中,随机事件是否发生是很重要的,但更重要的是事件发生的可能性的大小.它是随机事件的客观属性,是可以度量的.我们希望能将一个随机事件发生的可能性的大小用一个数来表达.为此,首先引入频率的概念,它描述了随机事件发生的频繁程度.

1.2.1 频率

定义 1.1 设在随机试验 E 中进行 n 次重复试验,其中事件 A 发生的次数 n_A 称为事件 A 发生的**频数**.称$\frac{n_A}{n}$为事件 A 发生的**频率**,记作 $f_n(A)$.

显然频率具有下述基本性质:

(1) $0\leqslant f_n(A)\leqslant 1$;

(2) $f_n(S)=1$;

(3) 若 $A_1,A_2,\cdots,A_k$ 是两两互不相容的事件,则 $f_n\left(\bigcup_{i=1}^{k}A_i\right)=\sum_{i=1}^{k}f_n(A_i)$.

以最简单的投硬币试验为例.投掷一枚均匀硬币,可能有两种结果:出现正面或反面.就一次试验而言,我们看不出这些结果发生的规律,但如果去做大量的试验,就可发现其中的一些规律性.历史上曾有人做过这类试验,结果见表 1-2.

表 1-2

试验者	试验次数	出现正面的次数	频率
德·摩根(De Morgan)	2048	1061	0.5181
蒲丰(Buffon)	4040	2048	0.5069
皮尔逊(K. Pearson)	12 000	6019	0.5016
皮尔逊(K. Pearson)	24 000	12 012	0.5005

从上面的试验数据可以看出,当试验次数相当大时,出现正面的频率稳定地在 0.5 左右摆动.一般而言,频率随试验次数的变化而变化,但当试验次数足够大时,频率又稳定地在某个常数附近摆动,这种"频率的稳定性"即是所谓的统计规律性,它揭示了隐藏在随机现象中的规律性.由随机事件的频率的稳定性可以看出,随机事件发生的可能性的大小是随机事件

本身固有的,可以用一个数来表示,于是我们可以给出随机事件概率的统计定义.

1.2.2 概率的定义

定义 1.2 设在相同条件下进行大量独立重复试验,若事件 A 的频率稳定地在某一确定值 p 附近摆动,则称此数值 p 为事件 A 发生的**概率**,记作 $P(A)$,此时 $P(A)=p$.

此定义给出了随机事件概率的近似计算方法.在许多实际问题中,当事件的概率不容易计算时,往往就可用频率近似代替概率,这正是1946年由冯·诺依曼(Von Neumann)和乌拉姆(Ulam)所建立的蒙特卡罗(Monte Carlo)方法的基本思想.

上述概率的统计定义有一定的局限性.柯尔莫哥洛夫(Kolmogorov)于1933年提出了概率的公理化结构,使概率论成为严谨的数学分支.下面给出概率的公理化定义.

定义 1.3 设 E 是随机试验,S 是它的样本空间,对于 E 中的事件 A,赋予一个实数 $P(A)$,若$P(A)$满足以下性质:

(1) 非负性:$P(A)\geqslant 0$;

(2) 规范性:$P(S)=1$;

(3) 可列可加性:若事件 $A_1,A_2,\cdots,A_n,\cdots$两两互不相容,有

$$P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty}P(A_i), \tag{1.1}$$

则称 $P(A)$为事件 A 发生的**概率**.

1.2.3 概率的性质

概率具有如下性质:

(1) $P(\varnothing)=0$. (1.2)

(2) 对于任何事件 A,有 $0\leqslant P(A)\leqslant 1$.

(3) 若事件 $A_1,A_2,\cdots,A_n$ 两两互不相容,则

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n}P(A_i). \tag{1.3}$$

式(1.3)称为概率的有限可加性.

(4) 设 A,B 是两个事件,若 $A\subset B$,则有

$$P(A)\leqslant P(B),$$
$$P(B-A)=P(B)-P(A). \tag{1.4}$$

(5) 对于任何事件 A,有

$$P(\bar{A})=1-P(A). \tag{1.5}$$

(6) 对于任何两事件 A,B,有

$$P(A\cup B)=P(A)+P(B)-P(AB). \tag{1.6}$$

式(1.6)可推广到 n 个事件的情形,即

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i\leqslant j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n). \tag{1.7}$$

1.2.4　古典概率模型

设试验 E 满足：

(1) 试验的样本空间所含有的模型可能结果(即样本点)只有有限个；

(2) 试验中每个基本事件发生的可能性相同.

具有上述两个特征的试验模型称为**等可能概型**. 由于这类随机现象是概率论发展初期的主要研究对象，故亦称**古典概率模型**，简称**古典概型**. 它在概率论中有很重要的地位：一方面，因为它比较简单，许多概念既直观而又容易理解；另一方面，它又概括了许多实际问题，有很广泛的应用. 利用概率的性质可以得出古典概型的计算公式.

对于随机试验 E，若样本空间 S 的样本点总数为 n，事件 A 所包含的样本点数为 m，则事件 A 发生的概率为

$$P(A)=\frac{\text{事件 }A\text{ 中包含的样本点数}}{\text{样本空间 }S\text{ 中样本点的总数}}=\frac{m}{n}. \tag{1.8}$$

式(1.8)也称为概率的古典定义.

例 1　将一枚硬币抛掷三次. 求：

(1)出现两次正面、一次反面的概率；(2)至少出现一次反面的概率.

解　记正面为 H，反面为 T，则相应的样本空间为

$$S=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}.$$

(1) 出现两次正面、一次反面的事件为

$$A=\{HHT,HTH,THH\}.$$

所以，概率为 $P(A)=\frac{3}{8}$.

(2) 至少出现一次反面的事件为

$$B=\{HHT,HTH,THH,HTT,HTH,THH,TTT\}.$$

所以，概率为 $P(B)=\frac{7}{8}$.

注：此处事件 B 发生的概率也可用其逆事件 $\bar{B}$ 来求，即

$$P(B)=1-P(\bar{B})=1-\frac{1}{8}=\frac{7}{8}.$$

例 2　设有 10 件产品，其中有 6 件正品，4 件次品，现从中任取 3 件. 求下列事件的概率：

(1) $A=\{$没有次品$\}$；　(2) $B=\{$只有 1 件次品$\}$；

(3) $C=\{$最多 1 件次品$\}$；　(4) $D=\{$至少 1 件次品$\}$.

解　因为产品无序，用组合计算. 从 10 件产品中任取 3 件的总的取法数为

$$n=\binom{10}{3}=120.$$

(1) 事件 A 没有次品，即取出的 3 件均为正品，且是从 6 件正品中取出的，所以

$$m_A=\binom{6}{3}=20,\quad P(A)=\frac{m_A}{n}=\frac{1}{6}.$$

(2) 事件 B 只有1件次品,应理解为有1件次品和2件正品,所以

$$m_B=\binom{4}{1}\binom{6}{2}=60,\quad P(B)=\frac{m_B}{n}=\frac{1}{2}.$$

(3) 事件 C 是事件 A 与事件 B 的和事件,且 $A\cap B=\varnothing$,所以

$$P(C)=P(A)+P(B)=\frac{1}{6}+\frac{1}{2}=\frac{2}{3}.$$

(4) 事件 D 可看成取出的3件产品分别有1件、2件、3件次品这三个事件之和,故

$$m_D=\binom{4}{1}\binom{6}{2}+\binom{4}{2}\binom{6}{1}+\binom{4}{3}\binom{6}{0}=100,\quad P(D)=\frac{m_D}{n}=\frac{5}{6}.$$

注:此处用逆事件来求更方便.即

$$P(D)=P(\overline{A})=1-P(A)=\frac{5}{6}.$$

例3 一口袋装有6只球,其中4只白球,2只红球.从袋中取球两次,每次随机地取一只.考虑两种取球方式:(a)第一次取一只球,观察其颜色后放回袋中,搅匀后再取一球.这种取球方式叫做**放回抽样**.(b)第一次取一球不放回袋中,第二次从剩余的球中再取一球.这种取球方式叫做**不放回抽样**.试分别就上面两种情况,求:(1)取到的两只球都是白球的概率;(2)取到的两只球颜色相同的概率;(3)取到的两只球中至少有一只是白球的概率.

解 (a) 放回抽样的情况.

以 A,B,C 分别表示事件"取到的两只球都是白球","取到的两只球都是红球","取到的两只球中至少有一只是白球".易知"取到的两只球颜色相同"这一事件即为 $A\cup B$,$C=\overline{B}$.

在袋中依次取两只球,每种取法为一基本事件,显然此时样本空间中仅包含有限个元素.且由对称性知,每个基本事件发生的可能性相同,因而可用古典概型来计算事件的概率.

因为是放回抽样,所以第一次和第二次都有6只球可供抽取,由乘法原理,共有 6×6 种取法,同理可处理事件 A 和 B 所对应的取法数.所以,可作如下计算.

(1) $P(A)=\frac{4\times4}{6\times6}=\frac{4}{9}$.

(2) $P(B)=\frac{2\times2}{6\times6}=\frac{1}{9}$,由于 $AB=\varnothing$,得

$$P(A\cup B)=P(A)+P(B)=\frac{5}{9}.$$

(3) $P(C)=P(\overline{B})=1-P(B)=\frac{8}{9}$.

(b) 不放回抽样的情况.

由读者自己完成.

例4 在1到2000的整数中随机地取一个整数.求:取到的整数能被6或8整除的概率.

解 设事件 A 为取到的整数能被6整除,事件 B 为取到的整数能被8整除,所求事件的概率为 $P(A\cup B)$.

由 $333<\frac{2000}{6}<334$,可得 $P(A)=\frac{333}{2000}$.

由$\frac{2000}{8}=250$,可得 $P(B)=\frac{250}{2000}$.

一个数能同时被 6 与 8 整除相当于能被 24 整除,又 $83<\frac{2000}{24}<84$,可得 $P(AB)=\frac{83}{2000}$,故所求概率为

$$p=P(A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{4}.$$

例 5 将 3 个小球随机地放入 4 个盒子中. 求:盒子中球的最多个数分别为 1,2,3 的概率.

解 3 个球放入 4 个盒子,是有重复的排列,总放法为 $n=4^3=64$.

(1) 每个盒子中球的最多个数为 1,即 3 个球分别放入 4 个盒子中的 3 个盒子,放法为 $m_1=\mathrm{P}_4^3=24$. 于是 $p_1=\frac{m_1}{n}=\frac{3}{8}$.

(2) 每个盒子中球的最多个数为 2,即 3 个球放入 4 个盒子中的两个盒子,放法为排列数 $\mathrm{P}_4^2=12$,其中一个盒子中有两个球,另一个盒子中有 1 个球,这 1 个球从 3 个球中取,取法为组合数 $\binom{3}{1}=3$ (或 2 个球从 3 个球中取,取法为组合数 $\binom{3}{2}=3$),所以球的放法为 $m_2=\mathrm{P}_4^2\binom{3}{1}=36$. 于是 $p_2=\frac{m_2}{n}=\frac{9}{16}$.

(3) 每个盒子中球的最多个数为 3,即 3 个球全放入 4 个盒子中的 1 个盒子里,放法为 $m_3=\mathrm{P}_4^1=4$. 于是 $p_3=\frac{m_3}{n}=\frac{1}{16}$.

注:本题中计算 p_2 时比较复杂,既要考虑盒子(用排列)又要考虑球(用组合),而利用概率的性质求解就简单多了.

先求出 p_1,p_3,因为 $p_1+p_2+p_3=1$,所以 $p_2=1-p_1-p_3=\frac{9}{16}$.

例 6 一年按 365 天计算. 现有 $k(k\leqslant 365)$个人聚会. 求:

(1)这 k 个人中至少有 2 人生日相同的概率;(2)这 k 个人中至少有 2 人生日在 1 月 1 日的概率.

解 求"至少……"的概率,这类问题从逆事件入手解决比较简单. k 个人生日(可重复排列)的排法总数为 $n=365^k$.

(1) 记 $A=\{$至少 2 人生日相同$\}$,则 $\bar{A}=\{k$ 个人生日都不相同(不重复排列)$\}$,于是 $m_{\bar{A}}=\mathrm{P}_{365}^k$. 所以有

$$P(A)=1-P(\bar{A})=1-\frac{\mathrm{P}_{365}^k}{365^k}.$$

(2) 记 $B=\{$至少 2 人生日在 1 月 1 日$\}$,$B_0=\{k$ 个人生日都不在 1 月 1 日$\}$,$B_1=\{k$ 个人中有 1 人生日在 1 月 1 日$\}$,则 $\bar{B}=B_0\cup B_1$. 而 B_1 又可看成$\{k$ 个人中有$(k-1)$个人生日不在1 月 1 日$\}$. 于是有

$$m_0=364^k,\quad m_1=\binom{k}{1}\times 364^{k-1}=k\times 364^{k-1},\quad m_{\bar{B}}=m_0+m_1,$$

$$P(\bar{B})=\frac{m_1+m_2}{n}=\frac{364^k+k\cdot 364^{k-1}}{365^k},\quad P(B)=1-P(\bar{B})=1-\frac{364^k+k\cdot 364^{k-1}}{365^k}.$$

注：问题(1)有如下有趣结果：

k	20	23	30	40	50	100
$P(A)$	0.411	0.507	0.706	0.891	0.970	0.999 999 7

1.3 条件概率与事件的独立性

1.3.1 条件概率

在实际当中会遇到求“事件 A 已发生的条件下事件 B 发生的概率”的问题，这就属于条件概率的问题.

定义 1.4 设 A,B 是两个事件，且 $P(A)>0$，称

$$P(B\mid A)=\frac{P(AB)}{P(A)} \tag{1.9}$$

为在事件 A 已发生的条件下事件 B 发生的**条件概率**.

注：由于增加了条件“事件 A 已发生”，所以 $P(B)$ 与 $P(B|A)$ 未必相等.

条件概率符合概率定义中的三个条件. 即

(1) $P(B|A)\geqslant 0$；

(2) $P(S|A)=1$；

(3) 若事件 $B_1,B_2,\cdots,B_n,\cdots$ 两两互斥，有

$$P\left(\bigcup_{i=1}^{\infty}B_i\,\middle|\,A\right)=\sum_{i=1}^{\infty}P(B_i\mid A).$$

从而对概率成立的结论也适用于条件概率. 例如，对任意的事件 A,B,B_1,B_2，有

$$P(\bar{B}\mid A)=1-P(B\mid A),$$

$$P(B_1\cup B_2\mid A)=P(B_1\mid A)+P(B_2\mid A)-P(B_1B_2\mid A).$$

例 1 某厂有职工 500 人，男女各半，男女职工中非熟练工人分别有 40 人与 10 人. 现从该厂职工中任意选取一人，试问：若已知选出的是女职工，她是非熟练工人的概率是多少？

解 记 $A=\{$选出的工人是女职工$\}$，$B=\{$选出的工人是非熟练工人$\}$，则 $AB=\{$选出的工人是非熟练女职工$\}$.

按式(1.9)得条件概率

$$P(B\mid A)=\frac{P(AB)}{P(A)}=\frac{10/500}{250/500}=\frac{1}{25}.$$

注：也可直接按条件概率的含义而不用按式(1.9)来求 $P(B|A)$. 既然已知选出的职工是女工，那么男职工就应当排除在考虑范围之外，因此“事件 A 发生条件下的事件 B”就相当于在全部女职工中任意选一人，并选出了非熟练工人，从而样本空间就是全部女职工，而上述事件所包含的样本点数就是女职工中的非熟练工人. 因此所求概率为

$$P(B\mid A)=\frac{10}{250}=\frac{1}{25}.$$

1.3.2 乘法公式

对任意两个事件 A,B,若 $P(A)>0$,由条件概率定义中的式(1.9)可得

$$P(AB)=P(A)P(B\mid A), \tag{1.10}$$

同理,当 $P(B)>0$ 时,则有

$$P(AB)=P(B)P(A\mid B). \tag{1.11}$$

上述两式称为概率的**乘法公式**.

我们还可以把两个事件的概率乘法公式推广到有限多个事件的积事件的情形.

设 $A_1,A_2,\cdots,A_n$ 为 n 个随机事件,且 $P(A_1A_2\cdots A_{n-1})>0$,则有

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2\mid A_1)\cdots P(A_{n-1}\mid A_1A_2\cdots A_{n-2})P(A_n\mid A_1A_2\cdots A_{n-1}). \tag{1.12}$$

例 2 某班共有 n 人,分得一张春节晚会的入场券.该班用抓阄的方式决定谁得到入场券,他们依次抓阄.求:

(1) 已知前 $k-1(k\leqslant n)$个人都没有抓到,第 k 个人抓到的概率;

(2) 第 k 个人抓到的概率.

解 设 A_i 表示第 i 个人抓到入场券,$i=1,2,\cdots,n$.

(1) $P(A_k\mid\overline{A}_1\overline{A}_2\cdots\overline{A}_{k-1})=\dfrac{1}{n-k+1}$.

(2)
$$\begin{aligned}P(A_k)&=P(\overline{A}_1\overline{A}_2\cdots\overline{A}_{k-1}A_k)=P(\overline{A}_1)P(\overline{A}_2\mid\overline{A}_1)\cdots\\&\quad P(\overline{A}_{k-1}\mid\overline{A}_1\overline{A}_2\cdots\overline{A}_{k-2})P(A_k\mid\overline{A}_1\overline{A}_2\cdots\overline{A}_{k-1})\\&=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\cdots\cdot\frac{n-k+1}{n-k+2}\cdot\frac{1}{n-k+1}=\frac{1}{n}.\end{aligned}$$

注:该题第二问也可用古典概型直接求出.此时为不放回抽样,考虑到每个人抓阄有先后次序,所以是排列问题(该问题也可作为组合问题来处理).k 个人抓阄,看成从 n 个元素中取 k 个元素的排列,有 P_n^k 种取法;第 k 个人抓到唯一的一张春节晚会的入场券,有 1 种取法;其余的 $k-1$ 个人抓阄,看成从剩下的 $n-1$ 个元素中取 $k-1$ 个元素的排列,有 P_{n-1}^{k-1} 种取法.利用乘法原理可得,第 k 个人抓到春节晚会的入场券这一事件共含 $1\cdot\mathrm{P}_{n-1}^{k-1}=\mathrm{P}_{n-1}^{k-1}$ 个样本点.所以有

$$P(A_k)=\frac{\mathrm{P}_{n-1}^{k-1}}{\mathrm{P}_n^k}=\frac{1}{n}.$$

第二问的结果表明,抓阄的结果与次序无关,体现出了抓阄的"公平性".

例 3 设一袋中装有 a 个红球,b 个白球,每次从袋中任取一球观察其颜色后放回,并同时放入 c 个与所取球同颜色的球.如果从袋中连续取球 4 次,试求第 1 次、第 2 次取到红球且第 3 次、第 4 次取到白球的概率.

解 设 A_i 表示第 i 次取到红球,$i=1,2,3,4$.于是有

$$\begin{aligned}P(A_1A_2\overline{A}_3\overline{A}_4)&=P(A_1)P(A_2\mid A_1)P(\overline{A}_3\mid A_1A_2)P(\overline{A}_4\mid A_1A_2\overline{A}_3)\\&=\frac{a}{a+b}\cdot\frac{a+c}{a+b+c}\cdot\frac{b}{a+b+2c}\cdot\frac{b+c}{a+b+3c}.\end{aligned}$$

1.3.3 事件的独立性

条件概率 $P(B|A)$反映了事件 A 的发生对事件 B 发生的影响.虽然 $P(B)$与 $P(B|A)$

不一定相等，但在某些情形下，$P(B)$与$P(B|A)$又是相等的，此时事件A发生与否对事件B的发生是没有影响的，事件A与事件B之间存在某种“独立性”.

例 4 某公司有行政管理人员100名，其中青年40名.该公司规定每天从所有行政人员中随机挑选一人为当天的值班人员，且不论其是否在前一天值过班.求：

(1) 已知第一天选出的是青年，第二天选出的也是青年的概率；

(2) 第二天选出的是青年的概率.

解 设A,B分别表示第一天、第二天选出的是青年，于是

(1) $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{\binom{40}{1}\binom{40}{1}}{\binom{100}{1}\binom{100}{1}}\Bigg/\dfrac{\binom{40}{1}}{\binom{100}{1}}=\dfrac{2}{5}$；

(2) $P(B)=\dfrac{40}{100}=\dfrac{2}{5}$.

实际上这一结果应该在意料之中，既然每天的候选人都是同样的这些行政人员，显然，前一天的选取结果对第二天的选取结果是不会产生影响的.

定义 1.5 设A,B是两个事件($P(A)P(B)>0$)，若

$$P(AB)=P(A)P(B), \tag{1.13}$$

则称事件A与事件B**相互独立**.此时有$P(B)=P(B|A)$.

注：①定义1.5在$P(A)=0$(或$P(B)=0$)的情形仍然成立；②若A与B相互独立，则$\overline{A}$与B,A与$\overline{B}$,$\overline{A}$与$\overline{B}$也分别相互独立.

可将独立性的概念进一步推广到多于两个事件的情形.

定义 1.6 对任意三个事件A,B,C，若

$$\begin{cases}P(AB)=P(A)P(B)\\P(AC)=P(A)P(C)\\P(BC)=P(B)P(C)\\P(ABC)=P(A)P(B)P(C)\end{cases},$$

则称事件A,B,C**相互独立**.

对任意n个事件$A_1,A_2,\cdots,A_n$，若

(1) $P(A_iA_j)=P(A_i)P(A_j)$, $1\leqslant i<j\leqslant n$；

(2) $P(A_iA_jA_k)=P(A_i)P(A_j)P(A_k)$, $1\leqslant i<j<k\leqslant n$；

$\vdots$

(3) $P(A_1A_2\cdots A_n)=\prod\limits_{i=1}^{n}P(A_i)$.

则称事件$A_1,A_2,\cdots,A_n$**相互独立**.

例 5 甲、乙、丙三人同时独立地向同一目标射击，他们射中目标的概率分别为0.4，0.5，0.7.求：

(1)至少有一人射中目标的概率；(2)恰有一人射中目标的概率.

解 设以A,B,C分别表示甲、乙、丙三人击中目标的事件.则有

(1) $P(A\cup B\cup C)=1-P(\overline{A}\overline{B}\overline{C})=1-P(\overline{A})P(\overline{B})P(\overline{C})$

$=1-0.6\times0.5\times0.3=0.91$;

(2) $P(A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C)=P(A\overline{B}\overline{C})+P(\overline{A}B\overline{C})+P(\overline{A}\overline{B}C)$

$=P(A)P(\overline{B})P(\overline{C})+P(\overline{A})P(B)P(\overline{C})+P(\overline{A})P(\overline{B})P(C)$

$=0.36.$

1.4 全概率公式和贝叶斯公式

1.4.1 全概率公式

1. 划分

定义 1.7 设 S 是随机试验 E 的样本空间,$B_1,B_2,\cdots,B_n$ 是试验 E 的 n 个事件,如果满足:

(1) $\bigcup\limits_{i=1}^{n}B_i=S$;

(2) 事件 $B_1,B_2,\cdots,B_n$ 两两互不相容,即 $B_iB_j=\varnothing,i\neq j,i,j=1,2,\cdots,n$.

则称 $B_1,B_2,\cdots,B_n$ 为样本空间 S 的一个**划分**,或是一个**完备事件组**.

2. 全概率公式

在计算一个复杂事件的概率时,先把复杂事件分解为若干个简单事件的和,计算出各简单事件的概率,然后再利用和事件概率的公式计算出复杂事件的概率,这是概率计算中常用的一种方法.

设 S 是试验 E 的样本空间,A 是试验 E 的任一事件,$B_1,B_2,\cdots,B_n$ 为 S 的一个划分.显然 $A=A\cap S=A\cap\left(\bigcup\limits_{i=1}^{n}B_i\right)=\bigcup\limits_{i=1}^{n}(AB_i)$,且 $AB_i\cap AB_j=\varnothing\,(i\neq j)$,于是 $P(A)=\sum\limits_{i=1}^{n}P(AB_i)$,又由式(1.11)得 $P(AB_i)=P(B_i)P(A|B_i)$.从而可得

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A\mid B_i). \tag{1.14}$$

式(1.14)称为**全概率公式**.其中,$P(B_i)P(A|B_i)$称为**部分概率**,$P(A)$称为**全概率**.全概率等于部分概率之和.

1.4.2 贝叶斯公式

利用条件概率的式(1.9)及全概率的式(1.14)可得

$$P(B_i\mid A)=\frac{P(B_iA)}{P(A)}=\frac{P(B_i)P(A\mid B_i)}{\sum\limits_{j=1}^{n}P(B_j)P(A\mid B_j)},\quad i=1,2,\cdots,n. \tag{1.15}$$

式(1.15)称为**贝叶斯**(Bayes)**公式**.

例 1 某产品由甲、乙、丙三个厂家供应.已知三个厂家的正品率分别为 0.95,0.90,0.80,三个厂家产品数所占比例为 2∶3∶5,混放在一起.(1)从中任取一件产品,求此件产品为正品的概率;(2)现取到一件产品为正品,则它是由甲、乙、丙三个厂中哪个厂生产的可能性大?

解 设事件 A={取到的产品为正品}, B_1={产品由甲厂生产}, B_2={产品由乙厂生产}, B_3={产品由丙厂生产}. 由已知条件得

$$P(B_1)=\frac{2}{10},\quad P(B_2)=\frac{3}{10},\quad P(B_3)=\frac{5}{10},$$

$$P(A\mid B_1)=0.95,\quad P(A\mid B_2)=0.90,\quad P(A\mid B_3)=0.80.$$

(1) 由全概率公式得

$$P(A)=\sum_{i=1}^{3}P(B_i)P(A\mid B_i)=\frac{2}{10}\times 0.95+\frac{3}{10}\times 0.90+\frac{5}{10}\times 0.80=0.86.$$

(2) 由贝叶斯公式得

$$P(B_1\mid A)=\frac{P(B_1)P(A\mid B_1)}{\sum\limits_{i=1}^{3}P(B_i)P(A\mid B_i)}=\frac{0.2\times 0.95}{0.86}=0.2209,$$

$$P(B_2\mid A)=\frac{P(B_2)P(A\mid B_2)}{\sum\limits_{i=1}^{3}P(B_i)P(A\mid B_i)}=\frac{0.3\times 0.90}{0.86}=0.3140,$$

$$P(B_3\mid A)=\frac{P(B_3)P(A\mid B_3)}{\sum\limits_{i=1}^{3}P(B_i)P(A\mid B_i)}=\frac{0.5\times 0.80}{0.86}=0.4651.$$

比较以上三个数据可知,该产品是由丙厂生产的可能性最大.

注: $P(B_i)$习惯上称为**先验概率**(先于试验的概率),它是根据以往经验确定的一种"主观概率",通常是已知的;而 $P(B_i|A)$称为 B_i 的**后验概率**(此概率在试验后才能确定),即在事件 A 发生之后再判断事件 B_i 发生的概率,可用贝叶斯公式求出.

例 2 设在某次世界女排赛中,中、日、美、古巴四队取得半决赛权;半决赛中,中国队对古巴队,日本队对美国队. 根据以往战绩,中国队战胜古巴队、日本队、美国队的概率分别为 0.5,0.8,0.6,日本战胜美国的概率为 0.5. 试求: 中国队最后夺得冠军的概率.

解 设中国队夺冠的事件为 A,中国队战胜古巴队的事件为 B,日本队战胜美国队的事件为 C. 由于中国队须战胜古巴队才能夺冠,所以有 $P(A)=P(AB)=P(A|B)P(B)$,记 $D=A|B$(事件 D 实际上表示中国队进入决赛并夺得冠军),由于决赛的对手可能是日本队,也可能是美国队,所以要用全概率公式. 由全概率公式得

$$P(D)=P(D\mid C)P(C)+P(D\mid \bar{C})P(\bar{C})=0.8\times 0.5+0.6\times 0.5=0.7.$$

再用乘法公式可得

$$P(A)=P(A\mid B)P(B)=P(D)P(B)=0.7\times 0.5=0.35.$$

小　　结

本章主要介绍了随机试验、样本空间、随机事件、古典概率模型、条件概率、划分等概率论的基本概念以及乘法公式、全概率公式、贝叶斯公式等常用概率计算公式,通过对上述知识的学习,使学生能运用相关公式求解相应的古典概型问题,培养学生的计算能力与分析、解决实际应用问题的能力.

基本要求：

(1) 了解随机试验、样本空间、随机事件的概念，掌握事件之间的关系与运算.

(2) 理解事件频率的概念，了解概率的定义，掌握概率的基本性质，能计算简单的古典概型问题.

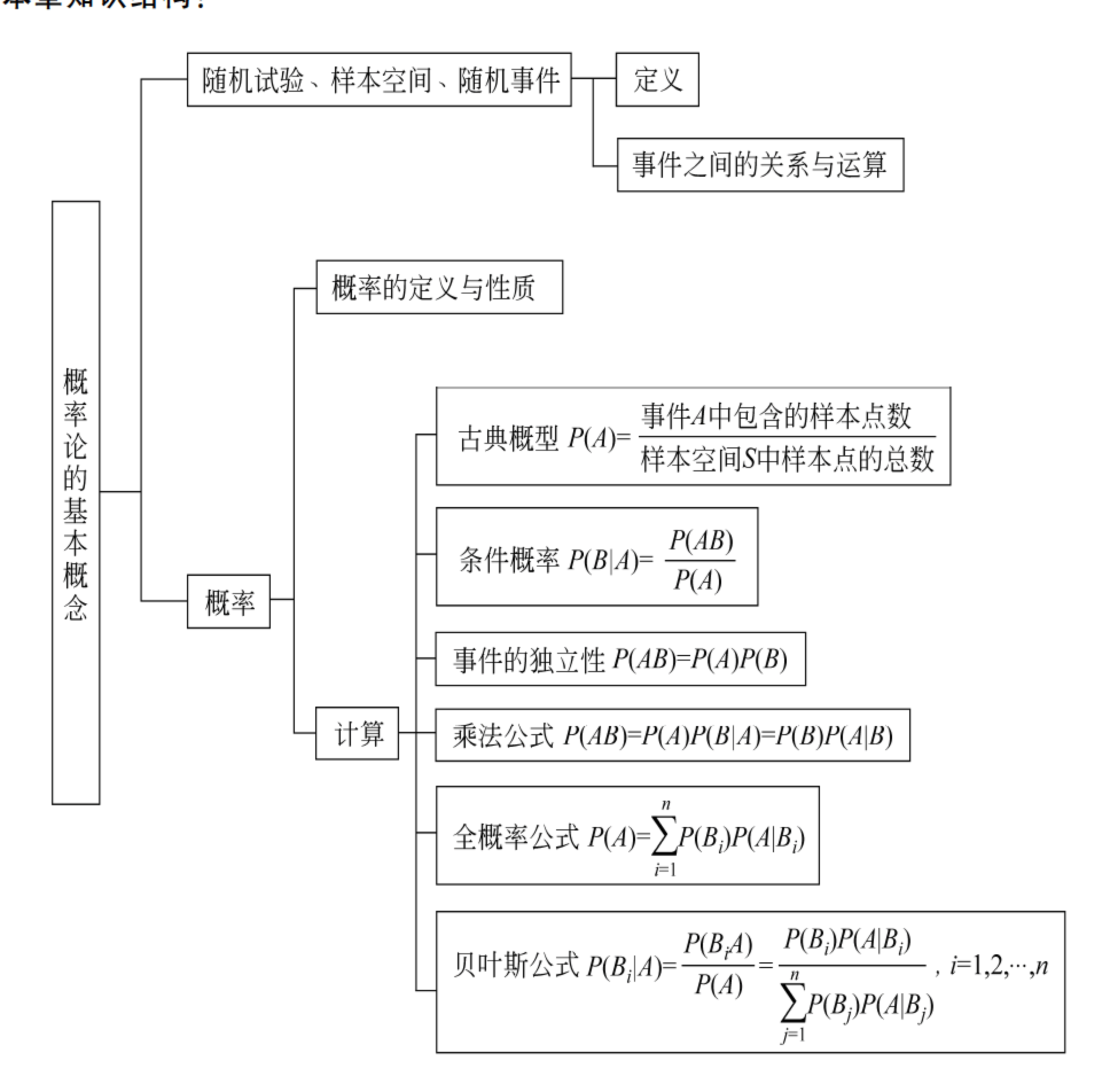
(3) 理解条件概率、事件的独立性的概念，掌握概率的乘法公式.

(4) 了解划分的概念，理解全概率公式和贝叶斯公式，并能利用其进行相应的计算.

重点：条件概率、乘法公式、全概率公式和贝叶斯公式、事件的独立性.

难点：运用条件概率、乘法公式、全概率公式和贝叶斯公式计算相关古典概型问题.

本章知识结构：



习　题　1

1. 写出下列随机试验的样本空间，并用样本点组成的集合表示给出的随机事件.

(1) 将一枚均匀硬币抛掷三次. $A=\{$第一次出现正面$\}$；$B=\{$两次出现同一面$\}$；$C=\{$至少有一次出现正面$\}$.

(2) 一个口袋中有 6 个外形完全相同的小球，编号分别为 1，2，3，4，5，6，从中任取 3 个球. $A=\{$球的最小编号为 2$\}$；$B=\{$球的编号全为奇数$\}$；$C=\{$球的编号之和小于 10$\}$.

(3) 掷两颗骰子. $A=\{$出现的点数之和为偶数且恰好有一次是 1 点$\}$；$B=\{$出现的点数

之和为奇数}.

2. 要在学院的学生中任选一名学生.设事件 $A=\{$选出的学生是男生$\}$;$B=\{$选出的学生是三年级学生$\}$;$C=\{$选出的学生是参加数学建模竞赛队员$\}$.则:

(1) 叙述事件 $AB\overline{C}$ 的含义.

(2) 在什么条件下,$ABC=C$ 成立?

(3) 什么时候关系式 $A\supseteq C$ 成立?

3. 某市发行 A,B,C 三种报纸.已知市民中订阅 A 报的有 45%,订阅 B 报的有 35%,订阅 C 报的有 30%,同时订阅 A 报及 B 报的有 10%,同时订阅 A 报及 C 报的有 8%,同时订阅 B 报及 C 报的有 5%,同时订阅三种报纸的有 3%.试求:

(1) 只订阅 A 报的概率 p_1;

(2) 同时只订阅 A 报及 B 报的概率 p_2;

(3) 恰好订阅两种报纸的概率 p_3;

(4) 至少订阅一种报纸的概率 p_4;

(5) 不订阅任何报纸的概率 p_5;

(6) 至多订阅一种报纸的概率 p_6.

4. 设 A,B 是两事件,且 $P(A)=0.5$,$P(B)=0.8$.问:(1)在什么条件下 $P(AB)$取到最大值,最大值是多少?(2)在什么条件下 $P(AB)$取到最小值,最小值又是多少?

5. 设事件 A,B,C 满足 $P(A)=P(B)=P(C)=0.2$,$P(AB)=P(BC)=0.1$,$P(AC)=0$.求事件 A,B,C 都不发生的概率.

6. 设 A,B 为两个随机事件,证明:

(1) $P(AB)=1-P(\overline{A})-P(\overline{B})+P(\overline{A}\,\overline{B})$;

(2) $1-P(\overline{A})-P(\overline{B})\leqslant P(AB)\leqslant P(A\cup B)\leqslant P(A)+P(B)$.

7. 将编号为 1,2,3,4 的四辆车按任意顺序停在车位上.问:四辆车自右向左或自左向右恰成 1,2,3,4 的顺序的概率是多少?

8. 从 $1,2,3,\cdots,9$ 这九个数字中任取一个,取后放回,共取五次.求下列事件的概率:

(1) 最后一次取出的数字是奇数;

(2) 五个数字全不相同;

(3) 2 恰好出现两次;

(4) 2 至少出现两次.

9. 从 52 张扑克牌中(不含大小王)任取 13 张牌.求:

(1) 其中有 4 张 A 的概率;

(2) 其中有 3 张黑桃、2 张红心、5 张方块、3 张草花的概率.

10. 某公司发出 20 件外形相同的货物,共有甲、乙、丙三种货物,其中甲 10 件,乙 6 件,丙 4 件.在运送途中所有标签脱落,交货人随意将这些货物发给顾客.问:一个订货 5 件甲、4 件乙和 2 件丙的顾客,能按其要求如数得到订货的概率是多少?

11. 将 8 本书随机地放到书架上,求其中 3 本数学书恰好放在一起的概率.

12. 在 1000 个产品中有 800 个正品、200 个次品.现任取 200 个产品.求:(1)恰好有 30 个次品的概率;(2)至少有 2 个次品的概率.

13. 已知 $P(A)=0.6$,$P(B)=0.4$,$P(A\overline{B})=0.5$,求 $P(B|A\cup\overline{B})$.

14. 假设一批产品中一、二、三等品各占60％，30％，10％，今从中随意取出一件，结果不是三等品.求取到的是一等品的概率.

15. 根据100多年来的气象记录知道，甲、乙两个相邻的城市一年中雨天的比例甲市占20％，乙市占14％，两地同时下雨占12％.试求：

(1) 甲市下雨的情况下，乙市出现雨天的概率；

(2) 甲市或乙市至少有一个城市下雨的概率.

16. 某设备具有甲、乙两套报警系统，单独使用时，甲、乙两套系统有效的概率分别为0.92和0.93；在甲系统失灵的条件下乙系统仍有效的概率为0.85.求：

(1) 发生意外时，这两套报警系统至少有一套有效的概率；

(2) 在乙系统失灵的条件下，甲系统仍有效的概率.

17. 甲、乙两人各自独立地对同一目标射击一次，设两人命中率分别为0.6和0.5.现已知目标被击中，求它是被甲击中的概率.

18. 电路由电池Ⅰ与两个并联的电池Ⅱ和Ⅲ串联而成.若电池Ⅰ，Ⅱ，Ⅲ损坏的概率分别是0.3，0.2，0.2，且各个电池的损坏与否是相互独立的，求发生断路的概率.

19. 某排球比赛的规则定为5盘3胜制.A，B两队的胜率分别为0.6和0.4，且每盘比赛胜否与之前比赛无关.若前两盘中A以2∶0领先，求最后A队获胜的概率.

20. 设$P(A)>0$，$P(B)>0$.证明：A，B相互独立$\Leftrightarrow P(A|B)=P(A|\bar{B})$.

21. 已知男人中有5％是色盲患者，女人中有0.25％是色盲患者.今从1000人的人群中(其中女性为600人)随机挑选一人，恰好是色盲患者，问此人是男性的概率.

22. 已知某产品的合格率为0.9.检验员在检验时，将合格品误认为次品的概率为0.01，而一个次品被误认为合格品的概率为0.05.求：(1)检查任一产品被认为是合格品的概率；(2)被认为合格品的产品实为次品的概率.

23. 某人到上海参加会议，他乘火车、轮船、汽车或飞机去的概率分别为0.5，0.2，0.2和0.1.如果他乘火车、轮船、汽车前去，迟到的概率分别为1/3，1/12和1/4，乘飞机不会迟到.如果他迟到了，求他是乘汽车去开会的概率.

24. 设有24只外形相同的球分装在三个盒子内，每盒装8只.第一盒内有5个标有2的球，3个标有3的球；第二盒内有红球和白球各4个；第三盒内有红球6个，白球2个.先从第一盒中取一球，若是“2”字球，则在第二盒中任取一球；若是“3”字球，则从第三盒中任取一球.求第二次取得的球是红球的概率.

25. 设有甲、乙两个袋子，甲袋中装有n只白球、m只红球；乙袋中装有N只白球、M只红球.今从甲袋中任取一只球放入乙袋中，再从乙袋中任意取一只球.问：(1)取到白球的概率是多少？(2)若从甲袋中任取两只球放入乙袋中，再从乙袋中任意取一只球，取到白球的概率又是多少？($n>1$，$m>1$)

26. 若有一传输信道，将三字母A，B，C分别输入信道，输出为原字母的概率为0.95，输出为其他字母的概率为0.025.现将3个字母串AAAA，BBBB，CCCC分别输入信道，输入的概率分别为0.3，0.4，0.3.设信道传输每个字母相互独立，已知输出字母串为ABCA.问：输入为字母串AAAA的概率是多少？

补充与提高

27. 从10双不同的鞋中任取6只，求下列事件的概率：

(1) 6只中至少有2只成一对；

(2) 6只中只有2只成一对；

(3) 6只中恰有4只成两对.

28. 随机地向半圆 $0<y<\sqrt{2ax-x^2}\ (a>0)$ 内掷一点，该点落在半圆内任何区域的概率与该区域的面积成正比. 求原点与该点的连线与 x 轴夹角小于 $\frac{\pi}{4}$ 的概率。

29. 甲、乙二人轮流投篮，规定由甲先开始，且甲每轮只投一次，而乙每轮连续投两次，先投中者为胜. 已知甲、乙每次投篮的命中率分别是 p 与 0.5. 求 p 为何值时，甲、乙获胜的概率相同.

30. 设 $0<P(A)<1, 0<P(B)<1$. 若 $P(A|B)+P(\bar{A}|\bar{B})=1$，试证明：$A$ 与 B 相互独立.

31. 某种仪器由三个部件组装而成，假设各部件质量互不影响且它们的优质品率分别为0.8，0.7与0.9. 已知如果三个部件都是优质品，则组装后的仪器一定合格；如果有一个部件不是优质品，则组装后仪器的不合格率为0.2；如果有两个部件不是优质品，则组装后仪器的不合格率为0.6；如果三个部件都不是优质品，则组装后仪器的不合格率为0.9. 求：(1)该组装仪器的不合格率；(2)如果已发现组装的仪器不合格，则它有几个部件不是优质品的概率最大.

32. 甲、乙两名选手进行比赛，已知每一局甲、乙获胜的概率的概率分别为0.6与0.4，比赛可采用3局2胜制或5局3胜制. 问在哪一种赛制下甲获胜的可能性大.

33. 用某种仪器检验电子元件，若元件是正品，经检验也是正品的概率为0.99；若元件是次品，经检验也是次品的概率为0.95. 当大批元件送来检验时，检验员只是随机地无放回地抽取3件，用该仪器对每1件进行独立检验. 若检验3件全是正品，则这批元件就可以出厂. 现送来元件100件，已知其中有4件次品，求这100件元件能出厂的概率.

34. 盒中放有12个乒乓球，其中有9个是新的. 第一次从其中任取3个来用，用后仍放回盒中. 第二次再从盒中任取3个，若已知第二次取出的球全是新球，求第一次取到的球全是新球的概率.

35. 假设一名射手在距目标250 m，200 m与150 m处进行射击的概率与射击距离成反比，并且在上述各距离射击命中目标的概率分别为0.05，0.1与0.2. 如果该射手在同一位置连续射击三次(假定每次击中与否互不影响)，现发现目标仅中一弹，求射手在各不同位置射击的概率.

36. 设有来自三个地区的各10名、15名和25名考生的报名表，其中女生的报名表分别为3份、7份和8份. 随机地取一个地区的报名表，从中先后抽出两份. 求：(1) 先抽到的一份是女生表的概率；(2) 已知后抽到的一份是男生表，求先抽到的一份是女生表的概率.

第 2 章

随机变量及其分布

2.1 随 机 变 量

在第 1 章中用样本空间的子集表示随机事件，这种表示方式对分析随机现象的统计规律有较大的局限性.为了全面地研究随机试验的结果，我们将随机试验的结果与实数对应起来，将随机试验的结果数量化，由此就产生了随机变量的概念.

我们已看到，有许多随机试验的结果与实数之间本身就存在着某种客观的联系.

例 1 袋中装有 5 个相同的球，它们的编号为 1,2,3,4,5，从中任取一个球，观察它的编号.

以 X 记所取球的编号，很明显 X 可能取值为 1,2,3,4,5，其样本空间为 $S=\{e\}=\{1,2,3,4,5\}$，则与每个实验结果相对应的 X 的取值为

$$X = X(e) = i,\quad e = i,\quad i = 1,2,3,4,5.$$

例 2 考察电话总机在单位时间内接到呼唤的次数.

记 X 表示单位时间内接到呼唤的次数，显然，X 可能取值为 $0,1,2,\cdots$，其样本空间为 $S=\{e\}=\{0,1,2,\cdots\}$，则

$$X = X(e) = k,\quad e = k,\quad k = 0,1,2,\cdots.$$

例 3 测试灯泡的寿命.

设 X 表示灯泡的寿命（以 h 计），X 的可能取值范围为 $[0,+\infty)$，其样本空间为 $S=\{e\}=\{t|t\geqslant 0\}$，则

$$X = X(e) = t,\quad e = t,\quad t \geqslant 0.$$

有些随机试验的结果与实数之间虽然没有上述那种自然的联系，但是可以人为地给它建立起一种对应关系.

例 4 观察上课时第一个进入教室的学生是男生（记为 B），还是女生（记为 G）.

样本空间为 $S=\{e\}=\{B,G\}$，为便于研究，将每一个结果用一个实数来代表，这样试验结果就数量化了.现在我们约定，当是男生时，取 $X=1$；是女生时，取 $X=0$.则有

$$X = X(e) = \begin{cases} 1, & e = B \\ 0, & e = G \end{cases}.$$

例 5 抛掷一枚硬币两次，观察出现正面（记为 H）还是反面（记为 T）.

样本空间为 $S=\{e\}=\{HH,HT,TH,TT\}$，设 X 表示硬币抛掷两次中出现正面的次数，则

$$X=X(e)=\begin{cases}0, & e=TT\\ 1, & e=HT \text{ 或 } TH.\\ 2, & e=HH\end{cases}$$

上面这些例子中出现了变量 X，这个用来描述随机试验的变量 X 取什么值，每次试验之前是不确定的. 试验之前只知道它所有可能的取值，因为它的取值依赖于试验的结果. 由于试验的结果是随机的，所以它的取值相应地也具有随机性，称这种变量为随机变量.

定义 21 设随机试验的样本空间为 $S=\{e\}$，若对于试验的每一个结果 $e\in S$，X 都有一个确定的实数 $X=X(e)$ 与之对应，则称 X 为**随机变量**.

在本书中，我们一般以大写字母 $X,Y,Z,\cdots$ 表示随机变量，而以小写字母 $x,y,z,\cdots$ 表示实数.

由定义可知，随机变量是定义在样本空间上的实值单值函数.

随机变量的取值随试验的结果而定，在试验之前不能预知它取什么值，且它的取值有一定的概率. 这些性质显示出随机变量与普通函数有着本质的差异.

引入随机变量后，随机事件就可以通过随机变量表达出来. 例如在例 2 中，事件“接到不少于 1 次呼唤”可用 $\{X\geqslant 1\}$ 来表示；“没有接到呼唤”可用 $\{X=0\}$ 表示. 又如在例 3 中，事件“寿命在 200 h 和 1000 h 之间”可用 $\{200\leqslant X\leqslant 1000\}$ 表示.

这样，对于随机事件的研究就可以转化为对随机变量的研究，这使我们有可能用微积分的方法对各种有关的问题进行深入的研究，使我们的研究更为方便.

按照随机变量可能取得的值，可以把它们分为两类：离散型随机变量与非离散型随机变量. 非离散型随机变量包括的范围很广，情况比较复杂，其中最重要也是实际中常遇到的是连续型随机变量. 因此，我们仅讨论离散型与连续型随机变量这两种基本类型的随机变量及其概率分布.

2.2 离散型随机变量及其概率分布

2.2.1 离散型随机变量及其分布律

定义 22 若随机变量所有可能取的值是有限个或可列无限个，则称这种类型的随机变量为**离散型随机变量**.

如 2.1 节例 1、例 4、例 5 中的随机变量，它们所有可能取的值是有限个，它们是离散型随机变量；又如 2.1 节例 2 中的随机变量，它所有可能取的值是可列无限个，它也是离散型随机变量. 而 2.1 节例 3 中的随机变量，它所可能取的值充满一个区间，是不能按一定次序一一列举出来的，因而它是非离散型随机变量. 本节只讨论离散型随机变量.

在研究一个随机试验的时候，我们不仅关心试验出现什么结果，更重要的是要知道出现这些结果的概率，即对随机变量，我们不仅要知道它取什么值，而且要知道它取这些值的概率. 也就是说，必须知道它的概率分布情况.

设离散型随机变量 X 所有可能取的值为 $x_1,x_2,\cdots,x_n,\cdots$，以及取这些值的概率为 $p_1,p_2,\cdots,p_n,\cdots$，即

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots. \tag{2.1}$$

则称式(2.1)为随机变量 X 的**概率分布律**,简称**分布律**. X 的分布律也可写成如下的表格形式:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

(2.2)

易知分布律具有如下两条性质:

(1) $p_k\geqslant 0, k=1,2,\cdots$;

(2) $\sum\limits_{k=1}^{\infty} p_k = 1$.

反过来,满足此两条性质的一组数一定是某个随机变量的分布律.

例 1 若离散型随机变量 X 的分布律为 $P\{X=k\}=\frac{c}{2^k}, k=1,2,\cdots$,试求常数 c.

解 由分布律的性质,有

$$\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} P\{X=k\} = \sum_{k=1}^{\infty}\frac{c}{2^k} = c\times\frac{\frac{1}{2}}{1-\frac{1}{2}} = c = 1.$$

此时 $p_k=\frac{1}{2^k}>0$,故 $c=1$ 为所求.

例 2 袋中有 2 个白球和 3 个黑球,每次从中任取一个球直到取得白球为止,求取球次数的分布律. 假定:

(1) 每次取出的黑球不再放回去;

(2) 每次取出的黑球仍放回去.

解 设随机变量 X 表示取球的次数.

(1) 每次取出的黑球不再放回,则 $X=1,2,3,4$,且

$$P\{X=1\}=\frac{2}{5}=0.4,$$

$$P\{X=2\}=\frac{3}{5}\times\frac{2}{4}=0.3,$$

$$P\{X=3\}=\frac{3}{5}\times\frac{2}{4}\times\frac{2}{3}=0.2,$$

$$P\{X=4\}=\frac{3}{5}\times\frac{2}{4}\times\frac{1}{3}\times\frac{2}{2}=0.1.$$

得所求分布律为

X	1	2	3	4
p_k	0.4	0.3	0.2	0.1

(2) 每次取出的黑球仍放回去,则 $X=1,2,3,\cdots$,且

$$P\{X=k\}=\left(\frac{3}{5}\right)^{k-1}\times\frac{2}{5}=0.4\times(0.6)^{k-1},\quad k=1,2,\cdots.$$

得所求分布律为

X	1	2	…	n	…
p_k	0.4	0.4×0.6	…	$0.4\times(0.6)^{n-1}$	…

由于在(2)中随机变量 X 取得它的可能值的概率恰为几何数列，所以这种分布叫做**几何分布**. 几何分布的一般情形是：$P\{X=k\}=(1-p)^{k-1}p, k=1,2,\cdots$，其中 $0<p<1$ 为常数. 根据几何级数的收敛性，易知

$$\sum_{k=1}^{\infty}P\{X=k\}=\sum_{k=1}^{\infty}(1-p)^{k-1}p=\frac{p}{1-(1-p)}=1.$$

求得了随机变量 X 的分布律后，我们不仅知道了 X 取每个可能值的概率，而且也可很容易地求出 X 取某个范围内的数值的概率. 如在例2(1)中：

$$P\{X<3\}=P\{X=1\}+P\{X=2\}=0.7,$$

$$P\{1<X<4\}=P\{X=2\}+P\{X=3\}=0.5.$$

下面介绍几种常用的离散型随机变量的分布.

2.2.2 离散型随机变量的常用分布

1. (0-1)分布

若随机变量 X 只可能取 0 和 1 这两个值，且它的分布律为

$$P\{X=k\}=p^k(1-p)^{1-k},\quad k=0,1,0<p<1, \tag{2.3}$$

即

X	0	1
p_k	$1-p$	p

(2.4)

则称 X 服从**(0-1)分布**或**两点分布**.

一般地，若一个随机试验只有两种可能的结果 A 和 $\overline{A}$，且 $P(A)=p(0<p<1)$，我们总能定义一个服从(0-1)分布的随机变量

$$X=\begin{cases}1, & A\text{ 发生}\\ 0, & \overline{A}\text{ 发生}\end{cases}$$

来描述这个随机试验的结果. 例如，对新婴儿的性别进行登记，检查产品的质量是否合格，抛掷一枚硬币，观察其正、反面出现情况等试验都可以用服从(0-1)分布的随机变量来描述.

2. 伯努利试验、二项分布

进行 n 次重复试验，在每次试验中只有两个可能结果 A 和 $\overline{A}$，假设每次试验的结果与其他各次试验的结果无关，事件 A 发生的概率 $P(A)$ 在每次试验中保持不变，则称这样的 n 次独立重复试验为 ***n* 重伯努利(Bernoulli)试验**. n 重伯努利试验是一种很重要的概率模型，它有广泛的应用.

例如：抛掷一枚硬币 n 次，观察其正、反面出现情况，就是 n 重伯努利试验. 又如抛一颗骰子 n 次，观察其是否出现"1"点，也是 n 重伯努利试验. 再如袋中装有 m 只白球，n 只红球，每次在袋中随机地取一只球观察其颜色，若连续取球 n 次并作放回抽样，就是 n 重伯努利试

验;然而若作不放回抽样,由于各次试验不再相互独立,因而就不再是 n 重伯努利试验了.

以 X 表示 n 重伯努利试验中事件 A 发生的次数,$p=P(A)$,$0<p<1$,X 是一个随机变量,我们来求它的分布律.

显然,X 所有可能取值为 $0,1,2,\cdots,n$.

由于各次试验相互独立,因此,事件 A 在指定的 $k(0\leqslant k\leqslant n)$次试验中发生,其他 $n-k$ 次试验中 $\overline{A}$ 发生(例如前 k 次 A 发生,而后 $n-k$ 次 $\overline{A}$ 发生)的概率为

$$P(\underbrace{A\cdots A}_{k}\underbrace{\overline{A}\cdots\overline{A}}_{n-k})=p^k(1-p)^{n-k}.$$

由于这种指定的方式有$\binom{n}{k}$种,且它们是两两不相容的事件,故在 n 次试验中事件 A 恰好发生 k 次的概率为

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k},\quad 记\quad q=1-p.$$

即有

$$P\{X=k\}=\binom{n}{k}p^kq^{n-k},\quad k=0,1,2,\cdots,n.$$

显然

$$P\{X=k\}\geqslant 0,\quad k=0,1,2,\cdots,n;$$

$$\sum_{k=0}^{n}P\{X=k\}=\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}=(p+q)^n=1.$$

一般地,有如下定义.

设随机变量 X 具有分布律

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k},\quad k=0,1,2,\cdots,n,\tag{2.5}$$

其中 $0<p<1$ 为常数,则称 X 服从以 n,p 为参数的**二项分布**,记为 $X\sim B(n,p)$. 这里 $\binom{n}{k}p^k(1-p)^{n-k}$恰好是二项式$(p+q)^n$ 展开式出现 p^k 的那项,故得名.

特别地,当 $n=1$ 时二项分布即为(0-1)分布.

例 3 已知 100 个产品中有 5 个次品,现从中有放回地取 3 次,每次任取 1 个,以 X 记所取 3 个产品中的次品数.(1)写出 X 的分布律;(2)求在所取的 3 个产品中恰有 2 个次品的概率.

解 将抽取一次产品看成是一次试验,这是三重伯努利试验. 依题意,每次试验取到次品的概率为 $p=5/100=0.05$,则 $X\sim B(3,0.05)$,于是可作如下求解.

(1) X 的分布律为

$$P\{X=k\}=\binom{3}{k}0.05^k\times(1-0.05)^{3-k},\quad k=0,1,2,3.$$

(2) $P\{X=2\}=\binom{3}{2}0.05^2\times 0.95=0.0071$.

若将本例中的“有放回”改为“无放回”,这就不是 n 重伯努利试验了,X 也就不服从二

项分布，而只能用古典概型来求解，方法为

$$P\{X=2\}=\frac{\binom{95}{1}\binom{5}{2}}{\binom{100}{3}}=0.0062.$$

这时 X 服从后面将要提及的超几何分布.

在实际问题中，真正在完全相同条件下进行试验是不多见的. 例如，向同一目标射击 n 次，这 n 次射击条件不可能完全一样，只是大致相同，可用伯努利概型来近似处理. 对于抽样问题来说，当原产品的批量相当大，而抽查的产品数量相对于原产品的总数来说又很小时，“无放回”可以当作“有放回”来处理. 这样做含有一些误差，但误差不大，于是可用式(2.5)来近似计算取到的产品中含有 k 个次品的概率.

例 4　某地块岩层上有一个 10 m 深的土层，石块随机分布在土层内. 建房时设计的桩群要打到岩层. 设土层可以分为 5 个独立层，每层深 2 m，打桩时每一个 2 m 层内碰到一块石头的概率为 0.1(碰到两块或更多块石头的概率忽略不计). 试求：

(1) 一根桩成功地打到岩层而未碰到任何石块的概率；

(2) 打到岩层时一根桩最多碰到一块石头的概率；

(3) 打到岩层时一根桩恰有两次碰到石块的概率；

(4) 一根桩一直打到第四层才第一次碰到石块的概率；

(5) 假设一座房屋的地基要求有一组 9 根这样的桩打到岩层，各桩打入情况相互独立，求打桩时不碰到石块的概率.

解　设打桩时碰到石块的层数为 X，则 $X\sim B(5,0.1)$，求解如下.

(1) $P\{X=0\}=\binom{5}{0}\times 0.1^0\times 0.9^5=0.5905.$

(2) $P\{X\leqslant 1\}=P\{X=0\}+P\{X=1\}=\binom{5}{0}\times 0.1^0\times 0.9^5+\binom{5}{1}\times 0.1^1\times 0.9^4$

$=0.9185.$

(3) $P\{X=2\}=\binom{5}{2}\times 0.1^2\times 0.9^3=0.0729.$

(4) 一根桩一直打到第四层才第一次碰到石块即前三层都未碰到石块，而第四层才首次碰到石块. 首次碰到石块的次数服从几何分布，故所求的概率为

$$0.9^3\times 0.1=0.0729.$$

(5) 题中所述即打桩时 9 根桩都未碰到石块，而各桩打入时又相互独立，故所求的概率为

$$0.5905^9=0.0087.$$

3. 泊松分布

设随机变量 X 的分布律为

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,2,\cdots,\tag{2.6}$$

其中 $\lambda>0$ 是常数，则称 X 服从以 λ 为参数的**泊松分布**，记为 $X\sim P(\lambda)$.

易知

$$P\{X=k\}\geqslant 0,\quad k=0,1,2,\cdots;$$

$$\sum_{k=0}^{\infty}P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda}=1.$$

泊松分布常见于社会生活与具有物理背景的问题中,如电话局在单位时间内收到用户的呼唤次数、车站在单位时间内到达的乘客数、放射物质在某段时间内放射的粒子数等.此外,基于下面的定理,泊松分布还可用来作二项分布的近似计算.

定理 21(泊松定理) 设随机变量 $X_n\sim B(n,p_n)$,且 $\lim\limits_{n\to\infty}np_n=\lambda$,其中 $\lambda>0$ 为常数,则有

$$\lim_{n\to\infty}P\{X_n=k\}=\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\cdots.$$

证明 记 $np_n=\lambda_n$,则

$$\begin{aligned}\binom{n}{k}p_n^k(1-p_n)^{n-k}&=\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}\\&=\frac{\lambda_n^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-k}.\end{aligned}$$

由于

$$\lim_{n\to\infty}\lambda_n^k=\lim_{n\to\infty}(np_n)^k=\lambda^k,$$

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)=1,$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{n-k}=\lim_{n\to\infty}\left[\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right]^{-\lambda_n}\left(1-\frac{\lambda_n}{n}\right)^{-k}=e^{-\lambda},$$

从而

$$\lim_{n\to\infty}P\{X=k\}=\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}.$$

该定理表明,若 $X\sim B(n,p)$,则当 n 比较大而 p 又很小时,有如下的泊松近似公式:

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\cdots,n;\lambda=np.$$

例 5 某人进行射击,设每次射击的命中率为 0.02,独立射击 400 次.试求至少命中两次的概率.

解 将一次射击看成是一次试验,设击中的次数为 X,则 $X\sim B(400,0.02)$,于是所求概率为

$$\begin{aligned}P\{X\geqslant 2\}&=1-P\{X=0\}-P\{X=1\}\\&=1-0.98^{400}-400\times 0.02\times(0.98)^{399}\\&=0.9972.\end{aligned}$$

直接计算的计算量较大,可利用泊松定理,因为 $n=400$ 足够大,$\lambda=np=8$ 不太大,所以查附表 A.1 可得所求概率为

$$P\{X\geqslant 2\}=1-P\{X=0\}-P\{X=1\}\approx 1-e^{-8}-8e^{-8}\approx 0.997.$$

例 5 的结果(概率很接近于 1)说明了以下两个事实.

(1) 虽然每次射击的命中率很小(为0.02),但如果射击400次,则击中目标至少两次几乎是可以肯定的.这一事实说明,一个事件尽管在一次试验中发生的概率很小,但在大量重复的独立试验中,这个事件的发生几乎是必然的.也就是说小概率事件,在一次试验中可以认为是不会发生的,但在大量重复的独立试验中是不可忽视的.

(2) 如果射手在400次射击中,击中目标的次数竟不到2次,由于 $P\{X<2\}\approx 0.003$ 很小,而这个小概率事件在一次试验中竟发生了,则根据**实际推断原理**(概率很小的事件在一次试验中几乎是不发生的),我们将有理由怀疑"每次射击的命中率为0.02"这一假设,即有理由认为该射手射击的命中率达不到0.02.

*4. 超几何分布

设一批同类产品共 N 件,其中有 M 件次品,从中任取 $n(n\leqslant N)$ 件,则这 n 件产品中的次品数 X 是一个离散型随机变量,其分布律为

$$P\{X=k\}=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\quad k=0,1,2,\cdots,\min\{M,n\}. \tag{2.7}$$

具有此分布律的随机变量,称为服从参数为 (N,M,n) 的**超几何分布**,记为 $H(N,M,n)$.

由上面的定义可知,超几何分布产生于无放回抽样.实际问题中经常会遇到服从超几何分布的随机变量,但当 N,M,n 较大时,上述的概率计算相当繁杂.注意到当 N 很大,n 较小时,上述的次品率 $p=M/N$ 在抽取前后的差异很小,进而可以证明 $N\to\infty$ 时超几何分布将趋于二项分布

$$\lim_{N\to\infty}\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}=\binom{n}{k}p^k(1-p)^{n-k},\quad p=\frac{M}{N}.$$

从而当 N 足够大而 n 不太大时,有如下的近似公式:

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\approx\binom{n}{p}p^k(1-p)^{n-k},\quad p=\frac{M}{N}.$$

再由泊松定理知,在一定的条件下超几何分布也可用泊松分布来近似计算.

2.3 随机变量的分布函数

由于随机变量的概率分布情况并非都能以其取某个值的概率表示,如非离散型随机变量 X,其可能取的值无法一一列举出来,因而就不能像离散型随机变量那样可以用分布律来描述它.在后面会看到,我们通常所遇到的非离散型随机变量取某个特定值的概率等于0.实际上,对于这类随机变量我们所关心的并不是它取某个特定值的概率.例如,在测量误差、元件的寿命的讨论中,我们感兴趣的是测量误差落在某个区间内的概率、寿命大于某个数的概率.因此,需要考虑的是随机变量的取值落在某个区间上的概率.而

$$P\{x_1 < X \leqslant x_2\} = P\{X \leqslant x_2\} - P\{X \leqslant x_1\},$$

这就告诉我们,要掌握随机变量 X 的统计规律,只要对任意的实数 x,知道 $P\{X\leqslant x\}$就够了.这一概率显然与 x 有关.下面引入随机变量分布函数的概念.

定义 23 设 X 是随机变量,x 是任一实数,则函数

$$F(x) = P\{X \leqslant x\}$$

称为随机变量 X 的**分布函数**.

由定义可知,$F(x)$是一个定义在$(-\infty,+\infty)$上、值域为$[0,1]$的函数,且对任意实数 $x_1,x_2(x_1<x_2)$,有

$$P\{x_1 < X \leqslant x_2\} = P\{X \leqslant x_2\} - P\{X \leqslant x_1\} = F(x_2) - F(x_1), \qquad (2.8)$$

因此,利用分布函数可以完整地表示随机变量的概率分布.

正是通过分布函数,我们将能用微积分的方法来研究随机变量.

如果将 X 看成数轴上的随机点,那么随机变量 X 的分布函数 $F(x)$的几何意义就是点 X 落在区间$(-\infty,x]$上的概率.

分布函数 $F(x)$具有以下性质:

(1) 单调非降性 若 $x_1<x_2$,则

$$F(x_1) \leqslant F(x_2).$$

(2) 有界性 $0\leqslant F(x)\leqslant 1$,且

$$F(-\infty) = \lim_{x\to-\infty} F(x) = 0, \quad F(+\infty) = \lim_{x\to+\infty} F(x) = 1.$$

(3) 右连续性

$$F(x+0) = \lim_{t\to x^+} F(t) = F(x).$$

以上的性质中,前两条由分布函数的定义直接可以得出,第三条的证明超出本书范围,从略.

反过来,任一满足这三条性质的函数,一定可以作为某一随机变量的分布函数.

有了随机变量 X 的分布函数,一些重要的随机事件的概率就可用分布函数 $F(x)$表示,如:

$$P\{X > a\} = 1 - P\{X \leqslant a\} = 1 - F(a),$$
$$P\{X < a\} = \lim_{x\to a^-} F(x) = F(a-0),$$
$$P\{a \leqslant X \leqslant b\} = F(b) - F(a) + P\{x = a\},$$
$$P\{X = a\} = F(a) - F(a-0).$$

例 1 已知随机变量 X 的分布函数为 $F(x) = \sum\limits_{0<n\leqslant x} \dfrac{1}{2^n}$,求 $P\{X<3\}$,$P\{X\geqslant 4\}$.

解

$$P\{X < 3\} = F(3-0) = \sum_{n=1}^{2} \frac{1}{2^n} = \frac{3}{4},$$

$$P\{X \geqslant 4\} = 1 - P\{X < 4\} = 1 - F(4-0) = 1 - \sum_{n=1}^{3} \frac{1}{2^n} = 1 - \frac{7}{8} = \frac{1}{8}.$$

对离散型随机变量,它的分布函数可根据概率分布律求出.

例 2 设随机变量 X 的分布律为

X	0	1	2	3
p_k	0.08	0.42	0.42	0.08

求 X 的分布函数,并求 $P\left\{X\leqslant\frac{1}{2}\right\}$,$P\left\{\frac{3}{2}<X\leqslant\frac{5}{2}\right\}$,$P\{2\leqslant X\leqslant 3\}$.

解

$$F(x)=P\{X\leqslant x\}=\sum_{x_k\leqslant x}P\{X=x_k\}=\begin{cases}0, & x<0\\0.08, & 0\leqslant x<1\\0.5, & 1\leqslant x<2,\\0.92, & 2\leqslant x<3\\1, & 3\leqslant x\end{cases}$$

$F(x)$是一分段函数,其图像如图 2-1 所示.

图 2-1

因此有

$$P\left\{X\leqslant\frac{1}{2}\right\}=F\left(\frac{1}{2}\right)=0.08,$$

$$P\left\{\frac{3}{2}<X\leqslant\frac{5}{2}\right\}=F\left(\frac{5}{2}\right)-F\left(\frac{3}{2}\right)=0.92-0.5=0.42,$$

$$P\{2\leqslant X\leqslant 3\}=F(3)-F(2)+P\{X=2\}=1-0.92+0.42=0.5.$$

可见离散型随机变量的分布函数 $F(x)$是一右连续的阶梯型函数,X 的每一可能取值 x_k 均为 $F(x)$的跳跃间断点,其跳跃度恰为 p_k,而在两个相邻跳跃点之间分布函数值保持不变,这一特征实际上是所有离散型随机变量的共同特征.反过来,如果一个随机变量 X 的分布函数 $F(x)$是阶梯型函数,则 X 一定是一个离散型随机变量,其概率分布律可由分布函数 $F(x)$唯一确定:$F(x)$的跳跃点全体构成 X 的所有可能取值,每一跳跃点处的跳跃高度则是 X 在相应点处的概率.

对于离散型随机变量来说,分布函数完全可以代替概率分布律的作用;不过,在表达和研究离散型随机变量的分布时,我们用得较多的还是概率分布律,因为它比较方便.

例 3 向半径为 2 m 的圆形靶射击,设随机变量 X 表示击中点与靶心 O 的距离.已知击中点落在以靶心 O 为中心、半径为 x 的圆内的概率与该圆的面积 πx^2 成正比,并且已知不会脱靶.试求 X 的分布函数.

解 由于不会发生脱靶的情况,所以 X 的一切可能取值都在区间$[0,2]$上,故可有以下结果.

当 $x<0$ 时,$F(x)=P\{X\leqslant x\}=0$.

当 $x\geqslant 2$ 时,$F(x)=P\{X\leqslant x\}=1$.

当 $0\leqslant x\leqslant 2$ 时,依题意,$P\{0\leqslant X\leqslant x\}=k\pi x^2$,其中 k 为常数.又$\{0\leqslant X\leqslant 2\}$是必然事件,即 $1=P\{0\leqslant X\leqslant 2\}=4k\pi$,故得 $k=\frac{1}{4\pi}$,即

$$P\{0\leqslant X\leqslant x\}=\frac{x^2}{4},$$

于是

$$F(x)=P\{X\leqslant x\}=P\{X<0\}+P\{0\leqslant X\leqslant x\}=\frac{x^2}{4}.$$

因此 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0\\ \dfrac{x^2}{4}, & 0\leqslant x<2.\\ 1, & 2\leqslant x\end{cases}$$

它的图形是一条连续曲线,如图 2-2 所示.



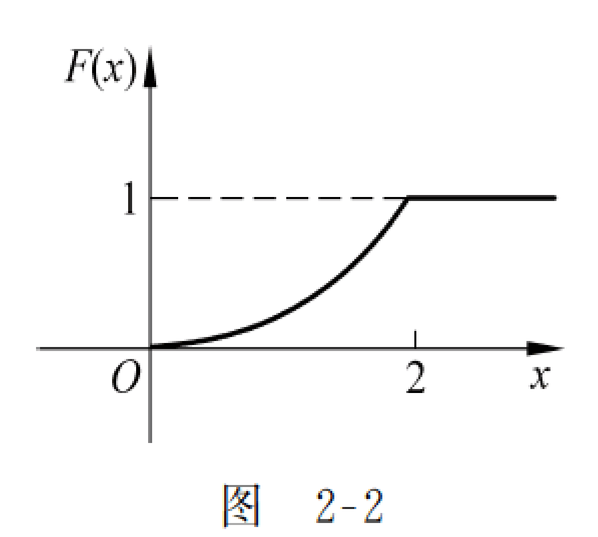


图 2-2

2.4 连续型随机变量及其概率分布

2.4.1 连续型随机变量及其概率密度

容易看到,在 2.3 节的例 3 中的分布函数 $F(x)$,对于任意 x 可表示为

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t,$$

其中

$$f(t)=\begin{cases}\dfrac{t}{2}, & 0<t<2\\ 0, & \text{其他}\end{cases}.$$

这就是说,$F(x)$恰是非负函数 $f(t)$在区间$(-\infty,x]$上的积分,在这种情况下我们称 X 为连续型随机变量.下面引入连续型随机变量的定义.

定义 24 若对于随机变量 X 的分布函数 $F(x)$,存在非负可积函数 $f(x)$,使对于任意实数 x 有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t, \tag{2.9}$$

则称 X 为**连续型随机变量**,称函数 $f(x)$为 X 的**概率密度函数**,简称**概率密度**或**密度函数**.

由定义知道,概率密度函数 $f(x)$具有下列性质:

(1) $f(x)\geqslant 0$.

(2) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$.

若一个函数满足上述两个性质,一定可以作为某一连续型随机变量的概率密度.

(3) 对于任意实数 $x_1,x_2(x_1<x_2)$,连续型随机变量 X 在$(x_1,x_2]$上取值的概率为

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(t)\mathrm{d}t. \tag{2.10}$$

由式(2.10)知道,连续型随机变量 X 落在区间$(x_1,x_2]$上的概率 $P\{x_1<X\leqslant x_2\}$等于区间$(x_1,x_2]$上曲线 $y=f(x)$之下的曲边梯形的面积,如图 2-3 所示.由性质(2)知道,介于曲线 $y=f(x)$与 Ox 轴之间的面积等于 1,如图 2-4 所示.

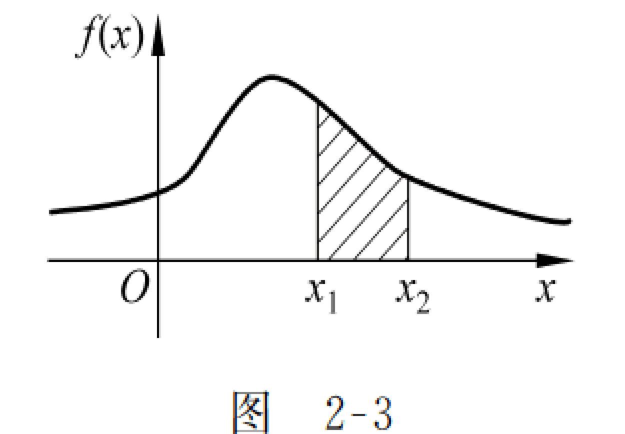


图 2-3

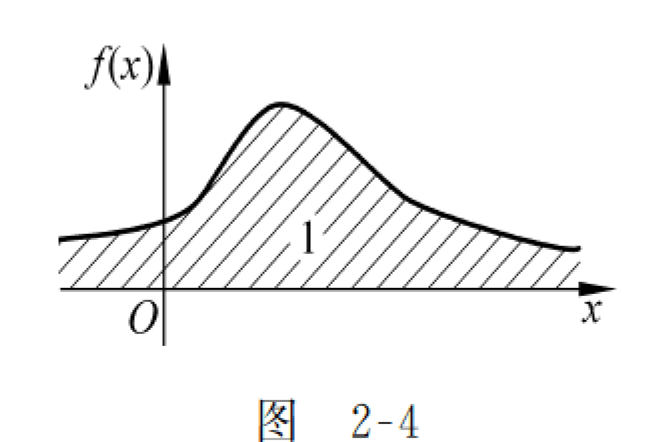


图 2-4

(4) 连续型随机变量的分布函数是连续函数,并且在 $f(x)$的连续点 x 处有

$$F'(x)=f(x).$$

由导数定义知

$$f(x)=\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{P\{x<X\leqslant x+\Delta x\}}{\Delta x},$$

故当 Δx 充分小时,有

$$P\{x<X\leqslant x+\Delta x\}\approx f(x)\Delta x.$$

即 $f(x)$不是概率,但 $f(x)$的取值确定了 X 在区间$(x,x+\Delta x]$上概率的大小,也就是说,$f(x)$的取值确定了 X 在点 x 附近的概率"疏密度",故称 $f(x)$为密度函数.

例 1 设随机变量 X 具有概率密度

$$f(x)=\begin{cases}kx, & 0\leqslant x<1\\ 1-\dfrac{x}{2}, & 1\leqslant x\leqslant 2.\\ 0, & \text{其他}\end{cases}$$

(1)确定常数 k;(2)求 X 的分布函数 $F(x)$;(3)求 $P\left\{\dfrac{1}{2}<X\leqslant 3\right\}$.

解 (1) 由性质(2)可得

$$1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_0^1 kx\,\mathrm{d}x+\int_1^2\left(1-\frac{x}{2}\right)\mathrm{d}x=\frac{k}{2}+\frac{1}{4},$$

解得 $k=\dfrac{3}{2}$,于是 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{3}{2}x, & 0\leqslant x<1\\ 1-\dfrac{x}{2}, & 1\leqslant x\leqslant 2.\\ 0, & \text{其他}\end{cases}$$

(2) X 的分布函数为

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\begin{cases}0, & x<0\\ \displaystyle\int_0^x\frac{3}{2}t\mathrm{d}t, & 0\leqslant x<1\\ \displaystyle\int_0^1\frac{3}{2}t\mathrm{d}t+\int_1^x\left(1-\frac{t}{2}\right)\mathrm{d}t, & 1\leqslant x<2\\ 1, & 2\leqslant x\end{cases}.$$

即

$$F(x)=\begin{cases}0, & x<0\\ \dfrac{3}{4}x^2, & 0\leqslant x<1\\ x-\dfrac{x^2}{4}, & 1\leqslant x<2\\ 1, & 2\leqslant x\end{cases}.$$

(3) $P\left\{\frac{1}{2}<X\leqslant 3\right\}=F(3)-F\left(\frac{1}{2}\right)=1-\frac{3}{4}\times\frac{1}{4}=\frac{13}{16}$,

或

$$P\left\{\frac{1}{2}<X\leqslant 3\right\}=\int_{\frac{1}{2}}^{3}f(t)\mathrm{d}t=\int_{\frac{1}{2}}^{1}\frac{3}{2}t\mathrm{d}t+\int_{1}^{2}\left(1-\frac{t}{2}\right)\mathrm{d}t=\frac{13}{16}.$$

需要指出的是,对于连续型随机变量 X 来说,它取任一指定实数值的概率均为 0 ,即 $P\{X=a\}=0$. 这是因为

$$\begin{aligned}P\{X=a\}&=\lim_{\Delta x\to 0^+}P\{a-\Delta x<X\leqslant a\}\\&=\lim_{\Delta x\to 0^+}(F(a)-F(a-\Delta x))=0.\end{aligned}$$

据此,在计算连续型随机变量落在某一区间的概率时,可以不必区分该区间是开区间或闭区间或半闭区间,即有

$$\begin{aligned}P\{a<X<b\}=P\{a\leqslant X\leqslant b\}&=P\{a<X\leqslant b\}=P\{a\leqslant X<b\}\\&=\int_a^b f(x)\mathrm{d}x.\end{aligned}$$

此外,这一结果也表明,概率为 0 的事件并不一定是不可能事件;同样概率为 1 的事件并不一定是必然事件.

例 2 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}A-B\mathrm{e}^{-2x}, & x\geqslant 0\\ 0, & x<0\end{cases}.$$

求:(1)常数 A,B;(2)概率密度函数 $f(x)$.

解 (1) 由分布函数的性质知

$$1=F(+\infty)=\lim_{x\to+\infty}F(x)=\lim_{x\to+\infty}(A-B\mathrm{e}^{-2x})=A.$$

又 $F(0)=A-B$,$F(0-0)=0$,$F(0+0)=A-B$,由 $F(x)$的连续性得 $A-B=0$,即 $B=1$,即有

$$F(x)=\begin{cases}1-\mathrm{e}^{-2x}, & x\geqslant 0\\ 0, & x<0\end{cases}.$$

(2) 显然

$$f(x)=F'(x)=\begin{cases}2\mathrm{e}^{-2x}, & x>0\\ 0, & x\leqslant 0\end{cases}.$$

这里 $F'_-(0)=0$,$F'_+(0)=2$,即 $F(x)$在 $x=0$ 处的导数是不存在的,但我们补充定义此点的密度为 $f(0)=0$(原则上,补充定义 $f(0)$为其他非负值也是可以的,因为改变密度函数个别点处的值不影响其在区间上的积分值).

对于连续型随机变量,在实用上和理论上还是使用 $f(x)$来描述较为方便.

2.4.2 连续型随机变量的常用分布

1. 均匀分布

设连续型随机变量 X 具有概率密度

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}, \tag{2.11}$$

则称 X 在区间 (a,b) 上服从**均匀分布**,记为 $X\sim U(a,b)$.

易知 $f(x)\geqslant 0$,且 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$.

容易求得均匀分布的分布函数

$$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\leqslant x<b.\\ 1, & b\leqslant x\end{cases}$$

$f(x)$ 及 $F(x)$ 的图形分别见图 2-5 和图 2-6.

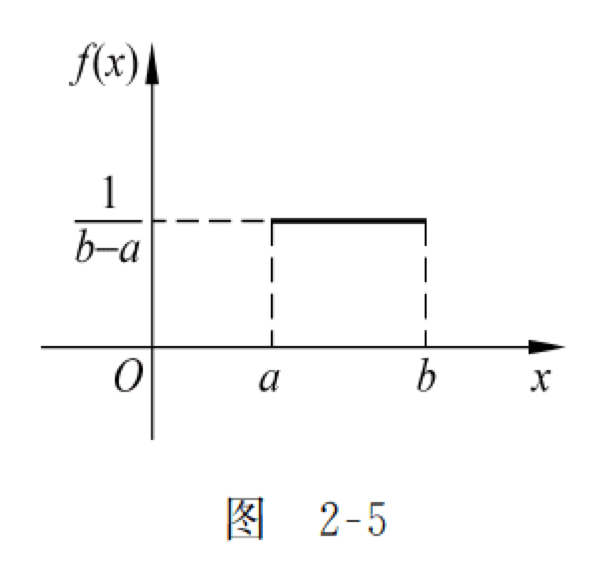


图 2-5

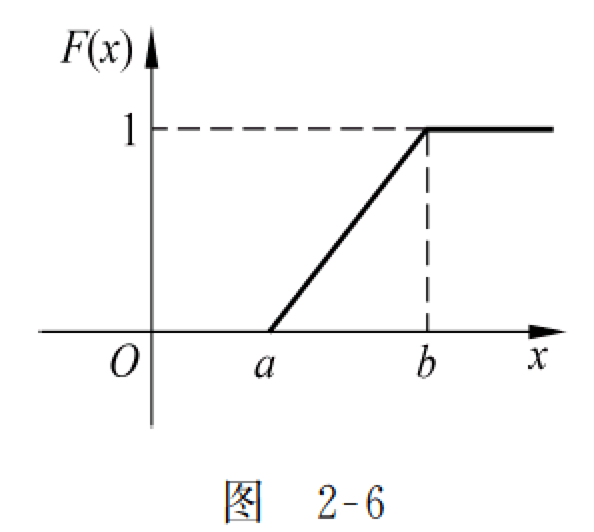


图 2-6

若 $X\sim U(a,b)$,则对于任一长度为 l 的子区间 $(c,c+l)$,$a\leqslant c<c+l\leqslant b$,有

$$P\{c<X\leqslant c+l\}=\int_c^{c+l}f(x)\mathrm{d}x=\int_c^{c+l}\frac{1}{b-a}\mathrm{d}x=\frac{l}{b-a}.$$

因此 X 的取值落在区间 (a,b) 内的任意子区间上的概率与子区间的长度成正比,而与子区间的位置无关,即 X 的取值落在任意长度相等的子区间上的概率"均匀"地分布在区间 (a,b) 内.

例 3 某长途汽车每天有两班,发车时间分别为 8:30 和 9:00,如果某乘客在 8:00～9:00 的任意时刻到达候车地点是等可能的,试求他等候不到 20 min 的概率.

解 设乘客于 8:00 过 X 分到达候车地点,则依题意,$X\sim U(0,60)$,其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{60}, & 0<x<60\\ 0, & \text{其他}\end{cases}.$$

故为使等候时间不超过 20 min,乘客必须且只需在 8:10～8:30 之间或 8:40～9:00 之间到达候车地点,因此所求概率为

$$P\{10<X\leqslant 30\}+P\{40<X\leqslant 60\}=\int_{10}^{30}\frac{1}{60}\mathrm{d}x+\int_{40}^{60}\frac{1}{60}\mathrm{d}x=\frac{2}{3}.$$

2. 指数分布

设连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases}\theta \mathrm{e}^{-\theta x}, & x>0\\ 0, & \text{其他}\end{cases}, \tag{2.12}$$

其中 $\theta>0$ 为常数,则称 X 服从参数为 θ 的**指数分布**,记为 $X\sim \mathrm{Exp}(\theta)$.

易知 $f(x)\geqslant 0$，且

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = \int_0^{+\infty} \theta \mathrm{e}^{-\theta x}\mathrm{d}x = -\mathrm{e}^{-\theta x}\Big|_0^{+\infty} = 1.$$

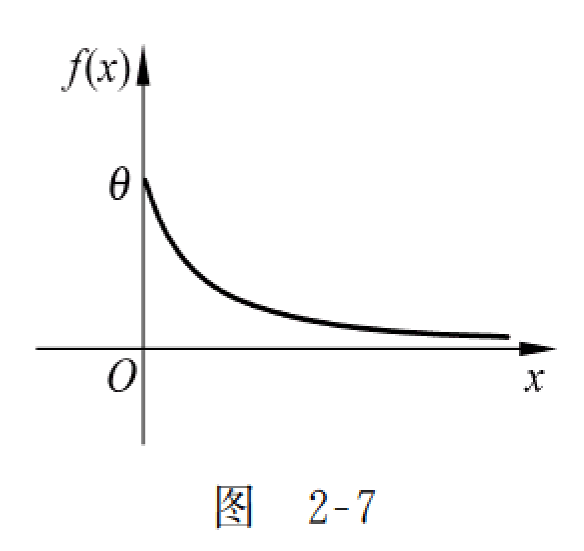


图 2-7

指数分布的密度曲线如图 2-7 所示，其相应的分布函数为

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \begin{cases} 1-\mathrm{e}^{-\theta x}, & x>0 \\ 0, & \text{其他} \end{cases}.$$

例 4 设打一次电话所用的时间(单位：min)服从参数 $\theta=\frac{1}{10}$ 的指数分布，若某人刚好在你前面走进公用电话亭. 求：(1)你将等待超过 10 min 的概率；(2)你将等待 5～10 min 才能走进此电话亭打电话的概率.

解 设 X 表示先进入那人打电话所用时间，则依题意，$X\sim \mathrm{Exp}\left(\frac{1}{10}\right)$，其概率密度为

$$f(x) = \begin{cases} 0.1\mathrm{e}^{-0.1x}, & x>0 \\ 0, & \text{其他} \end{cases}.$$

故所求概率为

(1) $P\{X>10\} = \int_{10}^{+\infty} 0.1\mathrm{e}^{-0.1x}\mathrm{d}x = -\mathrm{e}^{-0.1x}\Big|_{10}^{+\infty} = \mathrm{e}^{-1} \approx 0.368$；

(2) $P\{5<X<10\} = \int_5^{10} 0.1\mathrm{e}^{-0.1x}\mathrm{d}x = -\mathrm{e}^{-0.1x}\Big|_5^{10} = \mathrm{e}^{-0.5}-\mathrm{e}^{-1} \approx 0.239$.

在实践中，动植物及元件的寿命、电话问题中的通话时间、服务系统中的服务时间等通常都服从指数分布. 因此指数分布在排队论和可靠性理论等领域中有着广泛的应用.

指数分布的重要性还表现在它具有“无记忆性”，即对任意的 $s,t>0$，有

$$\begin{aligned} P\{X>s+t \mid X>s\} &= \frac{P\{X>s+t\}}{P\{X>s\}} = \frac{1-F(s+t)}{1-F(s)} = \frac{\mathrm{e}^{-\theta(s+t)}}{\mathrm{e}^{-\theta s}} = \mathrm{e}^{-\theta t} \\ &= P\{X>t\}. \end{aligned}$$

假如把 X 解释为某一元件的寿命，那么上式表明：已知元件已使用了 s 小时，则再使用 t 小时的概率与 s 无关. 这就是说，元件对它已使用过 s 小时没有记忆，所以有人戏称指数分布是“永远年轻的”.

3. 正态分布

设连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty<x<+\infty, \tag{2.13}$$

其中 $\mu,\sigma(\sigma>0)$ 为常数，则称 X 服从参数为 μ,σ 的**正态分布**，记为 $X\sim N(\mu,\sigma^2)$.

显然 $f(x)\geqslant 0$，现在来证明 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$.

令 $\frac{x-\mu}{\sigma}=t$，则

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t = \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1.$$

这里用到了高等数学中已知的结果：$\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t = \sqrt{2\pi}$.

正态分布的分布函数为(见图 2-8)

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt. \tag{2.14}$$

正态分布的概率密度 $f(x)$的图形如图 2-9 所示,它具有以下性质.

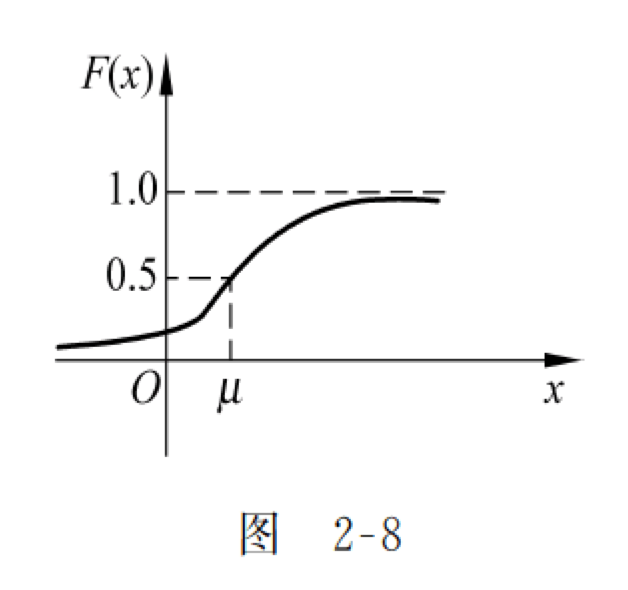


图 2-8

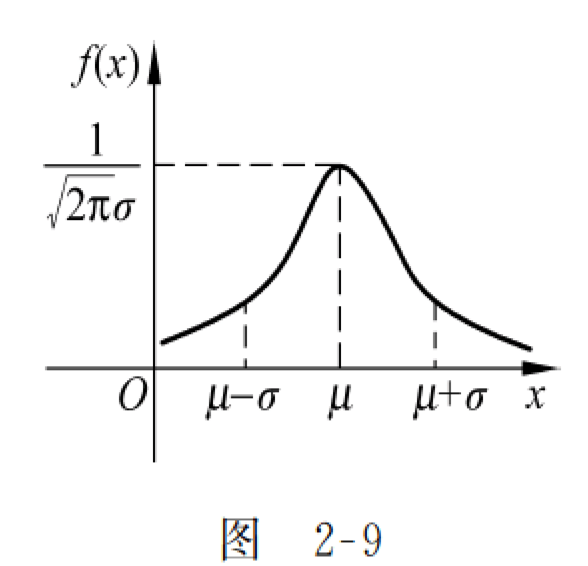
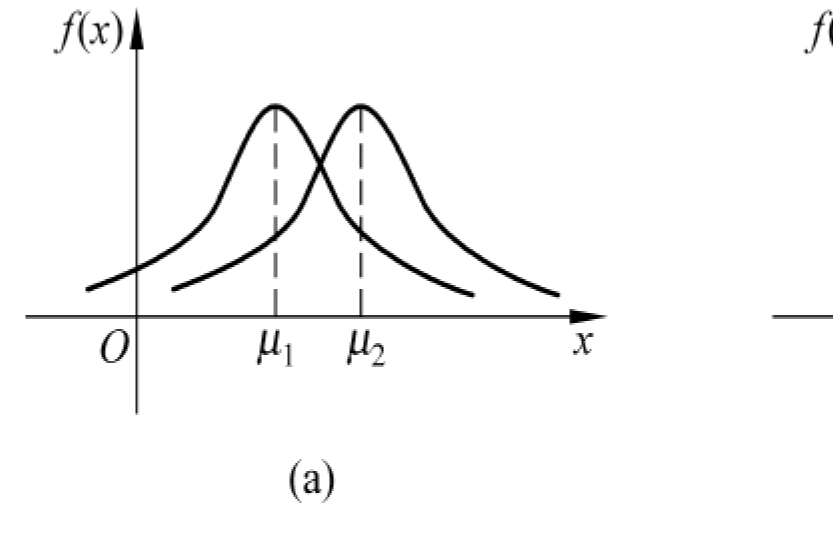


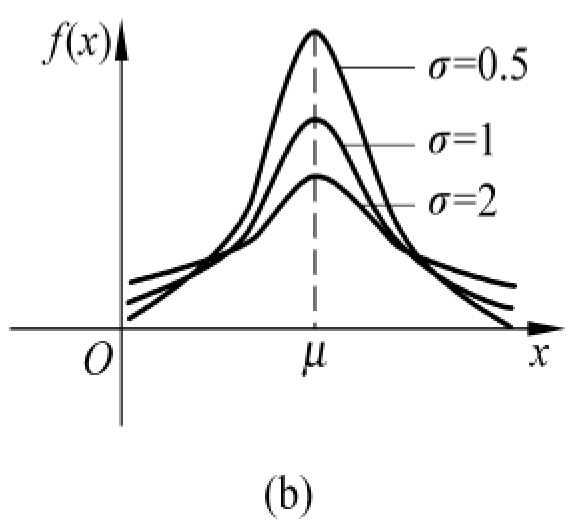
图 2-9

(1) 曲线关于 $x=\mu$ 对称,在 $x=\mu$ 处有最大值$\dfrac{1}{\sqrt{2\pi}\sigma}$,在 $x=\mu\pm\sigma$ 处有拐点,以 x 轴为水平渐近线.

(2) 若 σ 固定,改变 μ 的值,则曲线 $y=f(x)$沿 x 轴平行移动但不改变其形状,如图 2-10(a)所示,故参数 μ 确定了曲线 $y=f(x)$的位置;若 μ 固定,改变 σ 的值,则曲线的陡峭程度改变,但对称轴的位置不变,如图 2-10(b)所示,故参数 σ 确定了曲线 $y=f(x)$的形状.



(a)



(b)

图 2-10

正态分布是实际生活中最常见的一种分布.许多实际问题中的分布,都具有正态分布“中间大、两头小、左右对称”的特点,如人的身高、体重,学生的统考成绩,测量的误差,加工产品的尺寸,农作物的产量等,它们都服从或近似服从正态分布.后面还会看到,许多非正态分布的随机变量也和正态随机变量有着密切的联系,因此,正态分布是概率论中最重要的一种分布.

特别是,当 $\mu=0,\sigma=1$ 时,即 $X\sim N(0,1)$时,称 X 服从**标准正态分布**.其概率密度和分布函数分别用 $\varphi(x),\Phi(x)$表示,即有

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \tag{2.15}$$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt. \tag{2.16}$$

易知

$$\Phi(-x)=1-\Phi(x). \tag{2.17}$$

参见图 2-11.

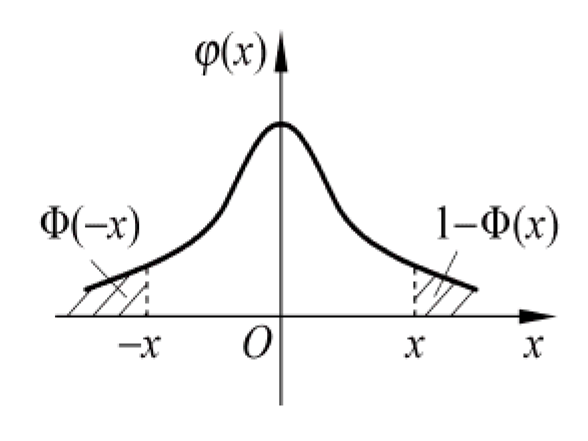


图 2-11

书后附录 A.2 中给出了 $\Phi(x)$ 的函数表.当 $x\geqslant 0$ 时,可在表中直接查到 $\Phi(x)$,而当 $x<0$ 时,可利用式(2.17)求得.

对于一般的正态分布 $N(\mu,\sigma^2)$,只要通过一个线性变换就能将它化成标准正态分布 $N(0,1)$.

引理 若 $X\sim N(\mu,\sigma^2)$,则 $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$.

证明 $Z=\frac{X-\mu}{\sigma}$ 的分布函数为

$$P\{Z\leqslant x\}=P\left\{\frac{X-\mu}{\sigma}\leqslant x\right\}=P\{X\leqslant \mu+\sigma x\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\mu+\sigma x}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t.$$

令 $\frac{t-\mu}{\sigma}=u$,得

$$P\{Z\leqslant x\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u=\Phi(x),$$

故

$$Z=\frac{X-\mu}{\sigma}\sim N(0,1).$$

于是,若 $X\sim N(\mu,\sigma^2)$,则

$$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right).\tag{2.18}$$

事实上,

$$F(x)=P\{X\leqslant x\}=P\left\{\frac{X-\mu}{\sigma}\leqslant\frac{x-\mu}{\sigma}\right\}=\Phi\left(\frac{x-\mu}{\sigma}\right),$$

由此可得,对于任意区间 $(x_1,x_2]$,有

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right).\tag{2.19}$$

式(2.19)表明:服从非标准正态分布的正态随机变量在一个区间内取值的概率可以用标准正态分布的分布函数来表示,而标准正态分布的分布函数 $\Phi(x)$ 的值可以查附表.

例 5 已知 $X\sim N(1,4)$,求 $P\{0<X\leqslant 1.6\}$.

解

$$\begin{aligned}P\{0<X\leqslant 1.6\}&=\Phi\left(\frac{1.6-1}{2}\right)-\Phi\left(\frac{0-1}{2}\right)=\Phi(0.3)-\Phi(-0.5)\\&=\Phi(0.3)-[1-\Phi(0.5)]=0.6179-1+0.6915=0.3094.\end{aligned}$$

例 6 已知 $X\sim N(\mu,\sigma^2)$,求 $P\{|X-\mu|<\sigma\}$,$P\{|X-\mu|<2\sigma\}$,$P\{|X-\mu|<3\sigma\}$.

解

$$\begin{aligned}P\{|X-\mu|<\sigma\}&=P\{\mu-\sigma<X<\mu+\sigma\}=\Phi(1)-\Phi(-1)\\&=2\Phi(1)-1=2\times 0.8413-1=0.6826,\end{aligned}$$

$$P\{|X-\mu|<2\sigma\}=2\Phi(2)-1=0.9544,$$

$$P\{|X-\mu|<3\sigma\}=2\Phi(3)-1=0.9974.$$

由例6可以看到，尽管正态变量的取值范围是$(-\infty,+\infty)$，但它的值落在$(\mu-3\sigma,\mu+3\sigma)$内几乎是肯定的事. 这就是通常所说的"$3\sigma$"原则(见图2-12).

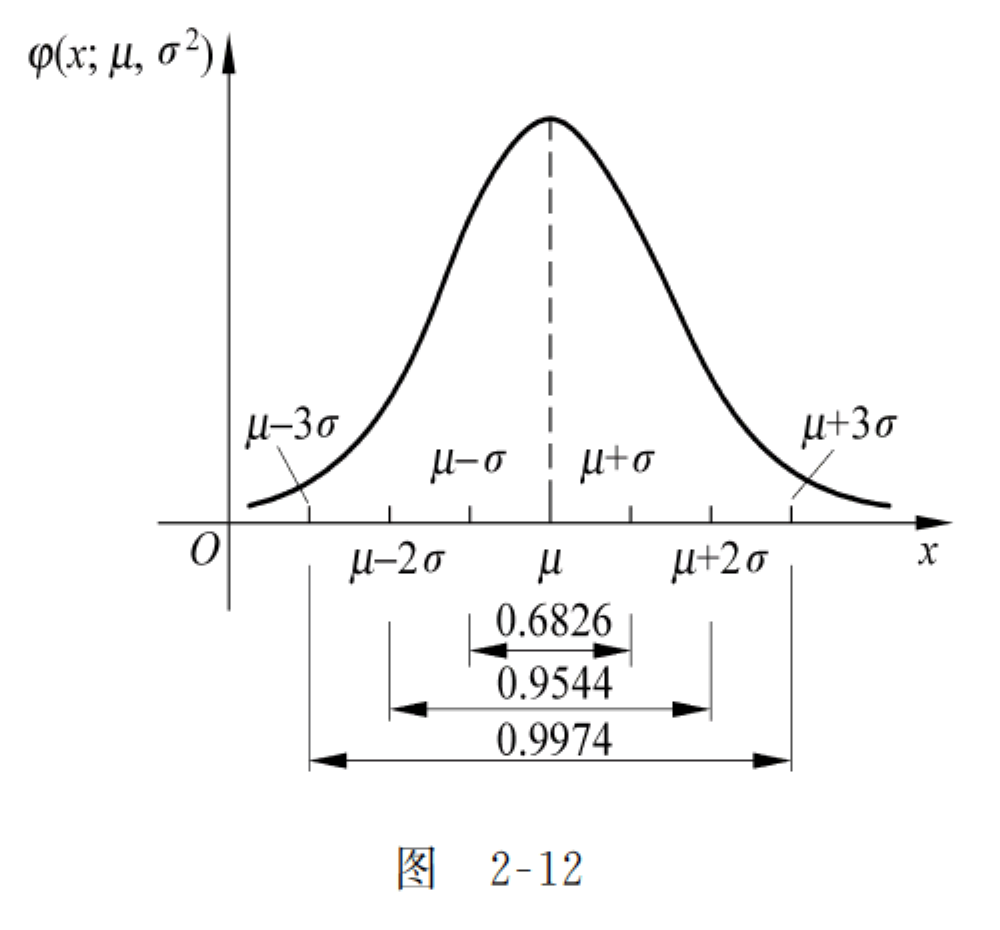


图 2-12

例7 在车床上加工金属圆杆，已知圆杆直径(单位：cm)$X\sim N(12.4,\sigma^2)$，规定直径在12.0～12.8 cm之间为合格品，要求产品合格的概率至少为0.95，试确定σ最多为多少.

解 依题意，需求σ，使得$P\{12.0<X\leqslant 12.8\}\geqslant 0.95$. 因为

$$P\{12.0<X\leqslant 12.8\}=\Phi\left(\frac{12.8-12.4}{\sigma}\right)-\Phi\left(\frac{12.0-12.4}{\sigma}\right)=2\Phi\left(\frac{0.4}{\sigma}\right)-1$$
$$\geqslant 0.95,$$

所以

$$\Phi\left(\frac{0.4}{\sigma}\right)\geqslant 0.975=\Phi(1.96).\quad \text{(反查标准正态分布表)}$$

又由$\Phi(x)$的单调性知

$$\frac{0.4}{\sigma}\geqslant 1.96,\quad \sigma\leqslant 0.204.$$

这就是说σ最多为0.204 cm.

为了今后在数理统计中应用方便，我们引进标准正态分布的上α分位点的概念.

定义25 设$X\sim N(0,1)$，若给定常数$\alpha(0<\alpha<1)$，存在数z_α满足

$$P\{X>z_\alpha\}=\alpha,$$

则称点z_α为标准正态分布的**上α分位点**(见图2-13).

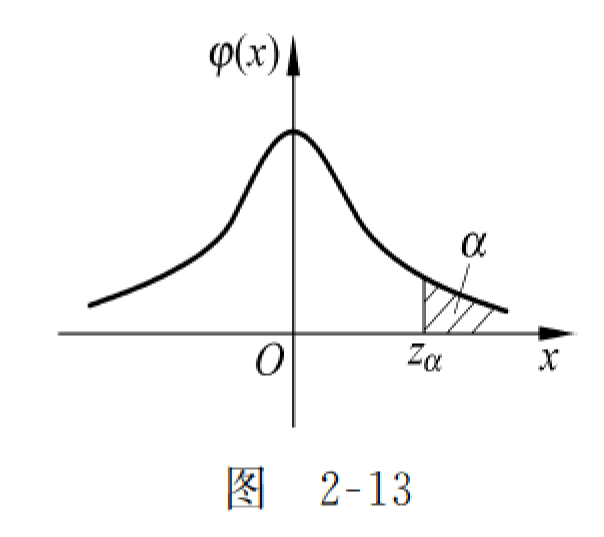


图 2-13

由关系式

$$\Phi(z_\alpha)=1-P\{X>z_\alpha\}=1-\alpha,\tag{2.20}$$

对于给定的α值，反查标准正态分布表可得z_α的值. 如$\alpha=0.05$，则由$\Phi(1.65)=1-0.05=0.95$，知$z_\alpha=1.65$.

另外，由$\varphi(x)$图形的对称性知道$z_{1-\alpha}=-z_\alpha$.

2.5 随机变量的函数的分布

在实际问题中，不仅要研究随机变量，而且往往还要研究随机变量的函数. 例如，电影院每放映一场电影所售出的票数是一个随机变量，而票房收入就是售出票数的函数，它当然也是一个随机变量. 有些情况下我们所关心的随机变量，它们的分布往往难以直接得到，但是与它们有关的另一些随机变量的分布却是容易知道的. 因此我们需要讨论，如何由已知的随机变量X的分布，去求得这个随机变量的函数$Y=g(X)$($g(x)$是已知的连续函数)的分布.

下面就两种情况分别进行讨论.

2.5.1 离散型随机变量的函数的分布

离散型随机变量 X 的函数 $Y=g(X)$ 仍然是一个离散型随机变量，它的分布律可直接由 X 的分布律求得.

若 X 的分布律为 $P\{X=x_i\}=p_i, i=1,2,\cdots$，则 $Y=g(X)$ 的全部不同的可能取值为 y_j，$j=1,2,\cdots$，其分布律为 $P\{Y=y_j\}=q_j, j=1,2,\cdots$，其中 q_j 是所有满足 $g(x_i)=y_j$ 的 x_i 对应的 X 的概率 $P\{X=x_i\}=p_i$ 的和，即

$$P\{Y=y_j\}=\sum_{g(x_i)=y_j} P\{X=x_i\}.$$

例 1 已知随机变量 X 的分布律为

X	0	1	2	3	4	5
p_k	0.10	0.14	0.16	0.18	0.20	0.22

求 $Y=(X-2)^2$ 的分布律.

解 Y 的全部不同的可能取值为 0,1,4,9，而

$$P\{Y=0\}=P\{(X-2)^2=0\}=P\{X=2\}=0.16,$$
$$P\{Y=1\}=P\{(X-2)^2=1\}=P\{X=1\}+P\{X=3\}=0.14+0.18=0.32,$$
$$P\{Y=4\}=P\{(X-2)^2=4\}=P\{X=0\}+P\{X=4\}=0.10+0.20=0.30,$$
$$P\{Y=9\}=P\{(X-2)^2=9\}=P\{X=5\}=0.22.$$

则 Y 的分布律为

X	0	1	4	9
p_k	0.16	0.32	0.30	0.22

2.5.2 连续型随机变量的函数的分布

若已知 X 的概率密度 $f_X(x)$ 或分布函数 $F_X(x)$，要求其函数 $Y=g(X)$ 的概率密度，通常是先求 Y 的分布函数 $F_Y(y)=P\{Y\leqslant y\}$，然后将 $F_Y(y)$ 关于 y 求导，就能得到 Y 的概率密度 $f_Y(y)$.

例 2 一食品厂一天的产量 X(单位：t)具有概率密度

$$f_X(x)=\begin{cases}2x, & 0<x<1\\ 0, & \text{其他}\end{cases},$$

一天的产值是 $Y=5X+1$(单位：千元)，求 Y 的概率密度 $f_Y(y)$.

解 现在先来求 $F_Y(y)$：

$$F_Y(y)=P\{Y\leqslant y\}=P\{5X+1\leqslant y\}=P\left\{X\leqslant \frac{y-1}{5}\right\}=F_X\left(\frac{y-1}{5}\right).$$

将上述等式两端对 y 求导，得 Y 的概率密度为

$$f_Y(y)=f_X\left(\frac{y-1}{5}\right)\frac{\mathrm{d}\left(\frac{y-1}{5}\right)}{\mathrm{d}y}=\frac{1}{5}f_X\left(\frac{y-1}{5}\right)$$

$$=\begin{cases}\dfrac{2}{5}\times\dfrac{y-1}{5}, & 0<\dfrac{y-1}{5}<1\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{2}{25}(y-1), & 1<y<6\\ 0, & \text{其他}\end{cases}.$$

例 3　设随机变量 X 具有概率密度 $f_X(x)$，$-\infty<x<+\infty$，求 $Y=X^2$ 的概率密度 $f_Y(y)$.

解　先来求 $F_Y(y)$：

由于 $Y=X^2\geqslant0$，因此，当 $y\leqslant0$ 时，$F_Y(y)=P\{Y\leqslant y\}=0$. 当 $y>0$ 时有

$$F_Y(y)=P\{Y\leqslant y\}=P\{X^2\leqslant y\}=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}$$
$$=F_X(\sqrt{y})-F_X(-\sqrt{y}).$$

将上述等式两端对 y 求导数，得 Y 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}[f_X(\sqrt{y})+f_X(-\sqrt{y})], & y>0\\ 0, & y\leqslant0\end{cases}.$$

例如，若 $X\sim N(0,1)$，即 X 的概率密度为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad -\infty<x<+\infty,$$

则 $Y=X^2$ 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}, & y>0\\ 0, & y\leqslant0\end{cases}.$$

在第 6 章中我们将会看到这是自由度为 1 的 χ^2 分布概率密度.

一般地，有如下便于应用的定理.

定理 22　设随机变量 X 具有概率密度 $f_X(x)$，$-\infty<x<+\infty$，$y=g(x)$在$(-\infty,+\infty)$内是严格单调的可导函数，则 $Y=g(X)$的概率密度为

$$f_Y(y)=\begin{cases}f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta\\ 0, & \text{其他}\end{cases}. \tag{2.21}$$

其中 $\alpha=\min\{g(-\infty),g(+\infty)\}$，$\beta=\max\{g(-\infty),g(+\infty)\}$，$h(y)$是 $g(x)$的反函数.

证明　当 $g(x)$是严格单调增加函数时，其反函数 $x=h(y)$存在，且也是严格单调增加，此时 $Y=g(X)$的可能取值落在区间$(\alpha=g(-\infty),\beta=g(+\infty))$内. 现在来求 Y 的分布函数 $F_Y(y)$.

当 $y\leqslant\alpha$ 时，$F_Y(y)=P\{Y\leqslant y\}=0$；

当 $\alpha<y<\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}=P\{X\leqslant h(y)\}=F_X[h(y)]$；

当 $y\geqslant\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=1$.

由此得 Y 的概率密度为

$$f_Y(y)=\begin{cases}f_X[h(y)]h'(y), & \alpha<y<\beta\\ 0, & \text{其他}\end{cases}.$$

若当 $g(x)$是严格单调减少函数时，其反函数 $x=h(y)$存在，且也是严格单调减少，此时 $Y=g(X)$的可能取值落在区间$(\alpha=g(+\infty),\beta=g(-\infty))$内. 现在来求 Y 的分布函数 $F_Y(y)$.

当 $y \leqslant \alpha$ 时，$F_Y(y)=P\{Y\leqslant y\}=0$；

当 $\alpha<y<\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}=P\{X\geqslant h(y)\}$
$=1-P\{X<h(y)\}=1-F_X[h(y)]$；

当 $y\geqslant\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=1$.

此时得 Y 的概率密度为

$$f_Y(y)=\begin{cases} f_X[h(y)][-h'(y)], & \alpha<y<\beta \\ 0, & \text{其他} \end{cases}.$$

综上所述得 Y 的概率密度为

$$f_Y(y)=\begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta \\ 0, & \text{其他} \end{cases}.$$

若 $f(x)$ 在有限区间 $[a,b]$ 以外等于零，则只需假设 $y=g(x)$ 在 $[a,b]$ 上是严格单调的可导函数，此时

$$\alpha=\min\{g(a),g(b)\},\quad \beta=\max\{g(a),g(b)\}.$$

例 4 设随机变量 $X\sim N(\mu,\sigma^2)$，概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<+\infty,$$

求 $Y=aX+b(a\neq0)$ 的概率密度 $f_Y(y)$.

解 因 $y=g(x)=ax+b(a\neq0)$ 严格单调，它的反函数为

$$x=h(y)=\frac{y-b}{a},\quad y\in(-\infty,+\infty),$$

且有

$$h'(y)=\frac{1}{a}.$$

由式(2.21)得 $Y=aX+b$ 的概率密度为

$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)=\frac{1}{\sqrt{2\pi}\,|a|\,\sigma}e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$
$$=\frac{1}{\sqrt{2\pi}\,|a|\,\sigma}e^{-\frac{[y-(b+a\mu)]^2}{2(a\sigma)^2}},\quad -\infty<y<+\infty.$$

由此可知，当 $X\sim N(\mu,\sigma^2)$ 且 $a\neq0$ 时，有

$$Y=aX+b\sim N(a\mu+b,(a\sigma)^2). \tag{2.22}$$

即服从正态分布的随机变量的线性组合仍服从正态分布.

特别地，在式(2.22)中令 $a=\frac{1}{\sigma},b=-\frac{\mu}{\sigma}$ 便可得到 $\frac{X-\mu}{\sigma}\sim N(0,1)$，这就是上一节引理的结果.

例 5 设 $X\sim U\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$，求 $Y=\sin X$ 的概率密度.

解 由题设知，X 的概率密度为

$$f_X(x)=\begin{cases} \frac{1}{\pi}, & -\frac{\pi}{2}<x<\frac{\pi}{2} \\ 0, & \text{其他} \end{cases}.$$

现在 $y=g(x)=\sin x$ 在 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 上严格单调增加，且有反函数 $x=h(y)=\arcsin y$，$y\in(-1,1)$，且 $h'(y)=\frac{1}{\sqrt{1-y^2}}$. 由式(2.21)得 $Y=\sin X$ 的概率密度为

$$f_Y(y)=\begin{cases}\frac{1}{\pi}\cdot\frac{1}{\sqrt{1-y^2}}, & -1<y<1\\ 0, & \text{其他}\end{cases}.$$

若在本例中 $X\sim U(0,\pi)$，由于此时 $y=g(x)=\sin x$ 在 $(0,\pi)$ 上不是单调函数，上述定理失效，应仍按一般的方法来做. 请同学们自行计算 $\left(f_Y(y)=\begin{cases}\frac{2}{\pi\sqrt{1-y^2}}, & 0<y<1\\ 0, & \text{其他}\end{cases}\right)$.

小　结

本章主要介绍了随机变量及其概率分布，主要内容有：随机变量、随机变量的分布函数、离散型随机变量的分布律、分布函数、连续型随机变量、概率密度函数、几个常用分布、随机变量的函数的概率分布.

基本要求：

(1) 理解随机变量、分布函数、离散型随机变量、连续型随机变量、分布律、概率密度函数的概念，及分布函数与概率密度函数的性质.

(2) 会用随机变量描述事件，会求随机变量的分布律，分布函数，概率密度函数.

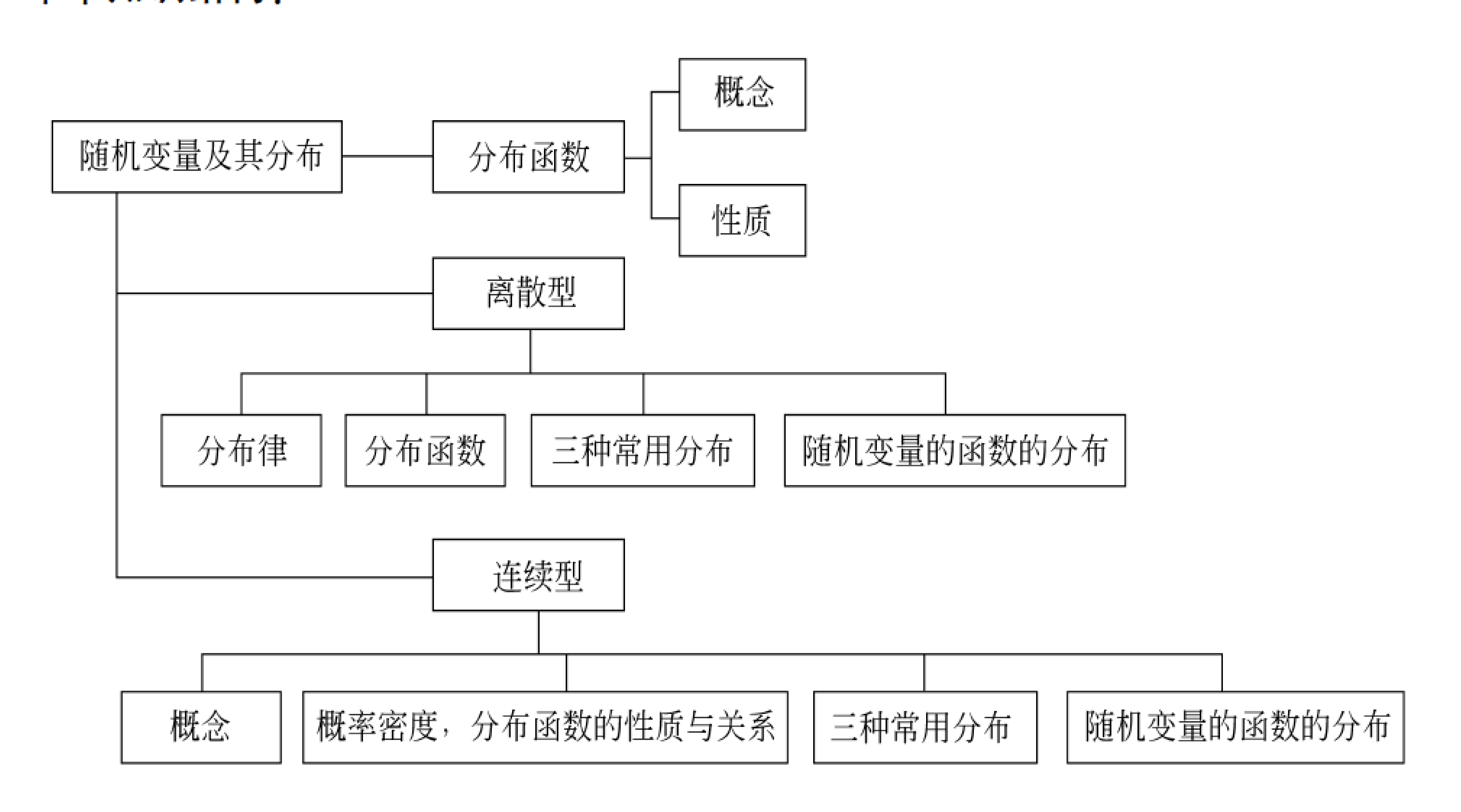
(3) 已知随机变量的分布函数，会求分布律或概率密度函数；已知随机变量的分布律或概率密度函数，会求分布函数，会求随机变量取值的概率.

(4) 熟练掌握几个常用分布，会求随机变量函数的分布律、分布函数，概率密度函数.

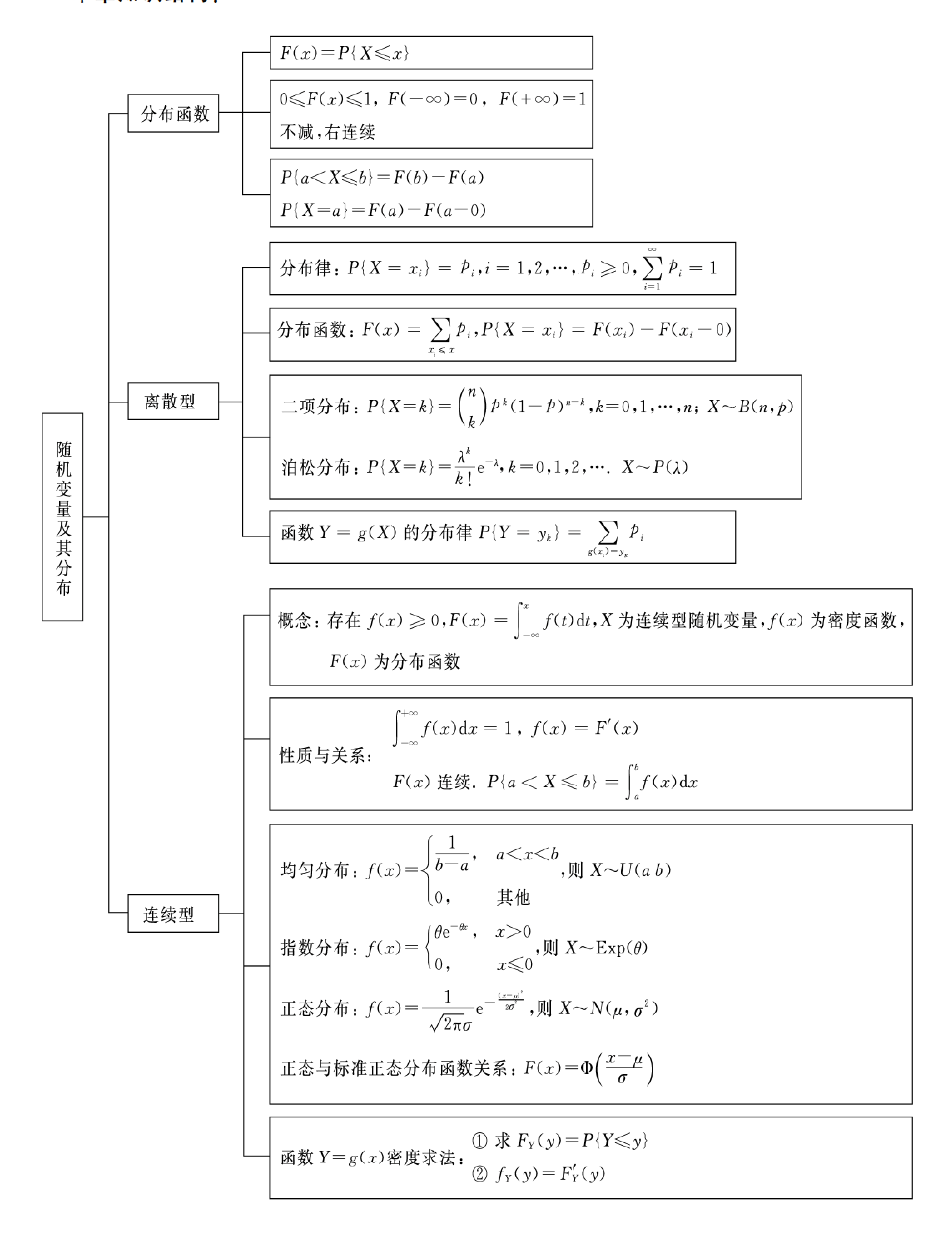
重点：随机变量的概念，分布函数与概率密度函数的概念、性质及计算，随机变量的函数的分布.

难点：求随机变量函数的概率密度函数.

本章知识结构：



本章知识结构：



习　题　2

1. 某路段有 3 个路口安置了红绿信号灯，且各路口亮什么颜色的灯是相互独立的. 红、绿颜色显示的时间为 1∶2，某人开车经过此路段，求此汽车首次遇到红灯前已通过的路口数 X 的分布律.

2. 一口袋中装有 5 个球，分别标有号码 1，2，3，4，5，现从这个口袋中任取 3 个球. 设

X 是取出球的号码中的最小值,写出随机变量 X 的分布律,并求 $P\{X>3\}$.

3. 一个人进行投篮,直到投中时为止.若这个人投中的概率为 0.4,求此人投篮次数 X 的分布律.

4. 设随机变量 X 的分布律为 $P\{X=k\}=\dfrac{c}{k+1}, k=0,1,3,5$. 试求:(1)常数 c;(2)$P\{X<3|X\neq 1\}$.

5. 设某产品的次品率为 0.1,检验员每天独立地检验 5 次.每次采用有放回抽样方式取 10 件检验.如果发现 10 件产品中有次品,就要调整设备,否则不需要调整.用 X 表示一天中调整设备的次数,求 X 的分布律.

6. 某地区有 5 个加油站,调查表明在任一时刻每个加油站被使用的概率为 0.2.求在同一时刻:(1)恰有两个加油站被使用的概率;(2)至少有 3 个加油站被使用的概率;(3)至多有 3 个加油站被使用的概率.

7. 一张答卷上有 5 道选择题,每道题列出了 4 个可能答案,其中有一个答案是正确的.某学生靠猜测能答对至少 3 道题的概率是多少?最可能对的题数是多少?

8. 设随机变量 X 服从泊松分布,且 $P\{X=1\}=P\{X=2\}$,求 $P\{X=5\}$.

9. (1)设 $X\sim$(0-1)分布,求 X 的分布函数,并做出其图形;(2)求第 1 题中的随机变量的分布函数,并求 $P\{X\leqslant 2\}$,$P\left\{\dfrac{1}{2}\leqslant X\leqslant 3\right\}$.

10. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0\\ \dfrac{x^2}{25}, & 0\leqslant x<5,\\ 1, & x\geqslant 5\end{cases}$$

求关于 t 的一元二次方程 $t^2+Xt+\dfrac{1}{4}(X+2)=0$ 有实根的概率.

11. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1\\ 0.3, & -1\leqslant x<1\\ 0.8, & 1\leqslant x<4\\ 1, & x\geqslant 4\end{cases},$$

求 X 的分布律.

12. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}A, & x<0\\ Bx^2, & 0\leqslant x<1\\ Cx-\dfrac{1}{2}x^2-1, & 1\leqslant x<2\\ 1, & x\geqslant 2\end{cases}.$$

求:(1)常数 A,B,C 的值;(2)X 的概率密度函数 $f(x)$;(3)$P\{1\leqslant X<4\}$.

13. 设连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}A\cos x, & -\dfrac{\pi}{2}\leqslant x\leqslant\dfrac{\pi}{2}.\\ 0, & \text{其他}\end{cases}$$

求：(1)常数 A；(2)X 的分布函数 $F(x)$，并画出 $F(x)$的图形.

14. 设某种元件的寿命 X(以 h 计)的概率密度为

$$f(x)=\begin{cases}\dfrac{100}{x^2}, & x>100\\ 0, & x\leqslant 100\end{cases}.$$

试问：(1) 若一个元件在使用 150 h 之后仍完好，那么该元件使用时间少于 200 h 的概率是多少？(2)若一个仪器中装有 3 个独立工作的这种元件，在使用 150 h 之后恰有一个元件损坏的概率是多少？

15. 设某公共汽车站每隔 8 min 就有一辆汽车通过，乘客在 8 min 内任一时刻到达汽车站是等可能的，求乘客候车时间不超过 5 min 的概率.

16. 修理某机器所需时间 X(单位：h)服从以 $\theta=\dfrac{1}{2}$为参数的指数分布. 试求：

(1) 修理时间超过 2 h 的概率；

(2) 若已持续修理了 8 h，总共需要至少 10 h 才能修好的概率.

17. 设 $X\sim N(3,2^2)$，解答以下问题：(1)求 $P\{2<X\leqslant5\}$，$P\{|X|>2\}$；(2)若 $P\{X>c\}=P\{X\leqslant c\}$，求 c；(3)要使 $P\{X>d\}\geqslant0.9$，问 d 至多是多少.

18. 设某地区成年男子的体重 X(以 kg 计)服从正态分布 $N(\mu,\sigma^2)$，已知 $P\{X\leqslant70\}=\dfrac{1}{2}$，$P\{X\leqslant60\}=\dfrac{1}{4}$.(1)确定 μ,σ；(2)若在这一地区随机地选出 5 名成年男子，求至少有 2 人体重超过 60 kg 的概率.

19. 设随机变量 X 的分布律为

X	-1	0	1	2
p_k	0.3	0.1	0.2	0.4

求 $Y=2X^2-1$ 的分布律.

20. 已知连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases}ax & 0<x<2\\ b-0.25x & 2\leqslant x\leqslant 4,\\ 0 & \text{其他}\end{cases}$$

且 $P\{1<X<3\}=\dfrac{3}{4}$.

求：(1)常数 a、b；(2)X 的分布函数；(3)$P\{X>3\}$；(4)$Y=-2X+1$ 的概率密度.

21. 设随机变量 X 在$(-1,1)$上服从均匀分布. 求以下随机变量的概率密度：

(1)$Y=3X+1$；(2)$Y=e^X$.

22. 设随机变量 $X\sim N(0,1)$，求 $Y=|X|$的概率密度.

补充与提高

23. 设离散型随机变量 X 的分布律为 $P\{X=k\}=b\lambda^k, k=1,2,\cdots$,且 $b>0$,则 λ 为(　　).

(A) 大于0的任意实数　　(B) $b+1$

(C) $\frac{1}{b+1}$　　(D) $\frac{1}{b-1}$

24. 若要 $f(x)=\cos x$ 可以成为随机变量 X 的概率密度,则 X 的可能取值区间为(　　).

(A) $\left[0,\frac{\pi}{2}\right]$　　(B) $\left[\frac{\pi}{2},\pi\right]$　　(C) $[0,\pi]$　　(D) $\left[\frac{3\pi}{2},\frac{7\pi}{4}\right]$

25. 一个盒子中装有4个红球、4个黑球和6个白球.现从中随机抽取3个球(不放回).假设抽出每个红球可以赢得1元,抽出每个黑球要输掉1元.以 X 表示赢得的总钱数.(1)写出随机变量 X 的分布律;(2)求我们赢钱的概率有多大.

26. 设随机变量 X 的密度函数为 $f(x)=Ae^{-|x|}, -\infty<x<+\infty$.试求:(1)常数 A;(2) X 的分布函数.

27. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\frac{1}{2}\cos\frac{x}{2}, & 0\leqslant x\leqslant\pi\\ 0, & \text{其他}\end{cases},$$

对 X 独立地重复观察4次,用 Y 表示观察值大于 $\frac{\pi}{3}$ 的次数,试求随机变量 Y 的概率分布.

28. 设随机变量 X 服从正态分布 $N(0,1)$,对给定的 $\alpha(0<\alpha<1)$,数 z_α 满足 $P\{X>z_\alpha\}=\alpha$.若 $P\{|X|<x\}=\alpha$,则 x 等于(　　).

(A) $z_{\frac{\alpha}{2}}$　　(B) $z_{1-\frac{\alpha}{2}}$　　(C) $z_{\frac{1-\alpha}{2}}$　　(D) $z_{1-\alpha}$

29. 设 X 服从区间 $[-1,9]$ 上的均匀分布,随机变量 Y 是 X 的函数:

$$Y=\begin{cases}-1, & X<1\\ 1, & X=1\\ 2, & 1<X\leqslant 6\\ 3, & 6<X\leqslant 9\end{cases},$$

求 Y 的概率分布.

30. 设连续型随机变量 X 有严格增加的分布函数 $F(x)$,试求 $Y=F(X)$ 的分布函数与密度函数.

31. 设有80台同类型设备,各台工作相互独立,发生故障的概率都是0.01,且一台设备的故障一个人能维修.考虑两种配备维护工人的方案:(1)由4个人维护,每人承包20台;(2)由3个人共同维护80台.试比较两种方案的优劣.

第 3 章

多维随机变量及其分布

在第 2 章中，我们所讨论的随机现象只涉及了一个随机变量，但在很多随机现象中，往往要涉及多个随机变量. 例如，炮弹弹着点的位置要用其横坐标 X 与纵坐标 Y 来确定；又如，健康体检时要检查身体的多项指标……，这都是多维随机变量问题. 多维随机变量的性质不仅与各个随机变量有关，而且还与它们之间的相互关系有关. 在本章中我们着重讨论二维随机变量.

3.1 二维随机变量的联合分布

3.1.1 二维随机变量及其分布函数

定义 3.1 设 E 是一个随机试验，$S=\{e\}$ 为它的样本空间，又设 $X=X(e)$，$Y=Y(e)$是定义在 S 上的随机变量，由它们组成的向量(X,Y)称为**二维随机变量**或**二维随机向量**.

与一维随机变量类似，可以用分布函数来讨论二维随机变量的概率分布. 下面引入二维随机变量(X,Y)的分布函数的定义.

定义 3.2 设(X,Y)是二维随机变量，对于任意实数 x,y，二元函数

$$F(x,y)=P\{(X\leqslant x)\cap(Y\leqslant y)\}\xlongequal{\text{def}}P\{X\leqslant x,Y\leqslant y\} \tag{3.1}$$

称为二维随机变量(X,Y)的**联合分布函数**，简称为(X,Y)的**分布函数**.

y
(x, y)
O
x

图 3-1

如果将二维随机变量(X,Y)看成平面上随机点的坐标，那么二维随机变量(X,Y)的分布函数 $F(x,y)$的几何意义就是随机点(X,Y)落在图 3-1 所示无穷矩形内的概率.

分布函数 $F(x,y)$具有下列性质：

(1) $0\leqslant F(x,y)\leqslant 1$，且

$F(-\infty,-\infty)=0$，　$F(+\infty,+\infty)=1$.

对于任意固定的 y，　$F(-\infty,y)=0$，

对于任意固定的 x，　$F(x,-\infty)=0$.

上面 4 个式子可以从几何上加以说明. 例如，在图 3-1 中无穷矩形的右面边界向左无限平移(即 $x\to-\infty$)，则随机点(X,Y)落在这个矩形内这一事件趋于不可能事件，故其概率趋于 0，即 $F(-\infty,y)=0$；又如当 $x\to+\infty$，$y\to+\infty$时，图 3-1 中的无穷矩形扩展到全平面，随机点(X,Y)落在其中这一事件趋于必然事件，故其概率趋于 1，即 $F(+\infty,+\infty)=1$.

(2) $F(x,y)$对 x,y 分别为单调不减函数.

(3) $F(x,y)$对 x,y 都是右连续函数.

(4) 对任意 $x_1\leqslant x_2,y_1\leqslant y_2$,有

$$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geqslant 0.$$

由分布函数的几何意义知上式为随机点(X,Y)落在图 3-2 中矩形内的概率,再由概率的非负性即得上式成立.

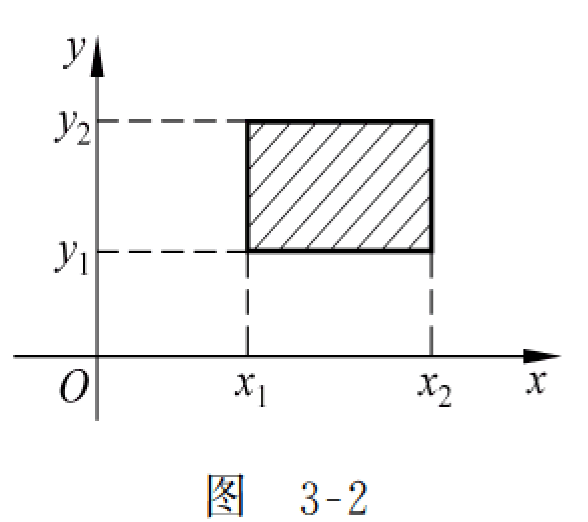


图 3-2

3.1.2 二维离散型随机变量

定义 3.3 设二维随机变量(X,Y)的所有可能取值为有限对或可列无限多对,则称(X,Y)为**二维离散型随机变量**.

定义 3.4 设二维离散型随机变量(X,Y)的所有可能取值为$(x_i,y_j)i,j=1,2,\cdots$,而

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots. \tag{3.2}$$

则称式(3.2)为随机变量(X,Y)的**联合分布律**,简称为(X,Y)的**分布律**.

上述(X,Y)的分布律也可写成如下表格的形式:

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$

显然,联合分布律满足 $p_{ij}\geqslant 0,i,j=1,2,\cdots$; $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1$.

例 1 一整数 X 等可能地在 1,2,3,4 中取值,另一整数 Y 等可能地在 $1\sim X$ 中取值,求(X,Y)的联合分布律.

解 X,Y 所有可能取的值均为 1,2,3,4.下面计算 $P\{X=i,Y=j\}$.

显然,当 $j>i$ 时,

$$P\{X=i,Y=j\}=0;$$

而当 $j\leqslant i$ 时,由概率乘法公式有

$$P\{X=i,Y=j\}=P\{X=i\}P\{Y=j\mid X=i\}=\frac{1}{4}\times\frac{1}{i},\quad i,j=1,2,3,4.$$

也可写成表格形式:

X \ Y	1	2	3	4
1	$\frac{1}{4}$	0	0	0
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

对于二维离散型随机变量(X,Y)，它的分布函数也可按下式求得：

$$F(x,y)=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y}p_{ij}.$$

这里，和式是对满足不等式 $x_i\leqslant x,y_j\leqslant y$ 的所有 i,j 求和.

3.1.3 二维连续型随机变量

定义 3.5 设(X,Y)为二维随机变量，$F(x,y)$为其联合分布函数. 如果存在非负函数 $f(x,y)$，使对任意的实数 x,y，都有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v, \tag{3.3}$$

则称(X,Y)为**二维连续型随机变量**，$f(x,y)$为(X,Y)的**联合概率密度**，简称**概率密度**.

联合概率密度 $f(x,y)$具有下列性质：

(1) $f(x,y)\geqslant 0$.

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1$.

(3) 若 G 是平面上某一区域，则有

$$P\{(X,Y)\in G\}=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y. \tag{3.4}$$

(4) 若 $f(x,y)$在点(x,y)连续，则有

$$\frac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y).$$

在几何上 $z=f(x,y)$表示空间的一个曲面，由性质(2)知，介于它和 xOy 平面之间的空间区域的体积为 1. 由性质(3)知，$P\{(X,Y)\in G\}$的值等于以 G 为底、以曲面 $z=f(x,y)$为顶的柱体体积.

例 2 设二维随机变量(X,Y)具有概率密度

$$f(x,y)=\begin{cases}k\mathrm{e}^{-(2x+y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases}.$$

(1)确定系数 k；(2)求分布函数 $F(x,y)$；(3)求概率 $P\{Y\leqslant X\}$.

解 (1) 由于 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1$，即

$$\int_0^{+\infty}\int_0^{+\infty}k\mathrm{e}^{-(2x+y)}\mathrm{d}x\mathrm{d}y=k\int_0^{+\infty}\mathrm{e}^{-2x}\mathrm{d}x\int_0^{+\infty}\mathrm{e}^{-y}\mathrm{d}y=k\times\frac{1}{2}\times 1=\frac{k}{2}=1,$$

因此可得

$$k=2.$$

(2)

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v=\begin{cases}\int_0^x\int_0^y 2\mathrm{e}^{-(2u+v)}\mathrm{d}u\mathrm{d}v, & x>0,y>0\\ 0, & \text{其他}\end{cases},$$

得

$$F(x,y)=\begin{cases}(1-\mathrm{e}^{-2x})(1-\mathrm{e}^{-y}), & x>0,y>0\\ 0, & \text{其他}\end{cases}.$$

(3) 将 (X,Y) 看作是平面上随机点的坐标，即有 $\{Y\leqslant X\}=\{(X,Y)\in G\}$，其中 G 为 xOy 平面直线 $y=x$ 下方部分，如图 3-3 所示.

于是

$$P\{Y\leqslant X\}=P\{(X,Y)\in G\}=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^{+\infty}\int_y^{+\infty}2\mathrm{e}^{-(2x+y)}\mathrm{d}x\mathrm{d}y=\frac{1}{3}.$$

图　3-3

3.1.4　两个常用的分布

1. 均匀分布

设 D 为 xOy 平面上的有界区域，其面积为 A. 若二维随机变量 (X,Y) 的联合概率密度

$$f(x,y)=\begin{cases}\dfrac{1}{A}, & (x,y)\in D\\ 0, & \text{其他}\end{cases}, \tag{3.5}$$

则称 (X,Y) 服从区域 D 上的**二维均匀分布**.

若 D_1 是区域 D 的子区域，其面积为 A_1，则

$$P\{(X,Y)\in D_1\}=\iint_{D_1}f(x,y)\mathrm{d}x\mathrm{d}y=\frac{1}{A}\iint_{D_1}\mathrm{d}x\mathrm{d}y=\frac{A_1}{A}.$$

可见此概率与 D_1 在 D 内的位置无关，仅与 D_1 的面积有关，这就是均匀分布中的"均匀"的含义.

2. 二维正态分布

若二维随机变量 (X,Y) 的联合概率密度为

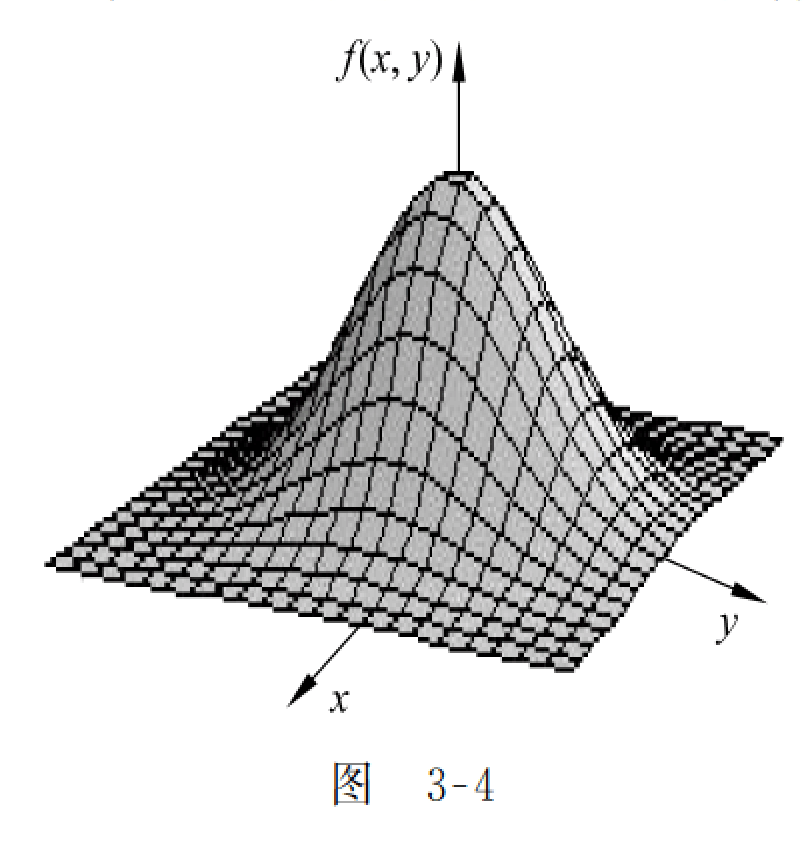


图　3-4

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]},$$

$$-\infty<x<+\infty,\quad -\infty<y<+\infty, \tag{3.6}$$

其中 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 均为常数，且 $\sigma_1>0,\sigma_2>0,|\rho|<1$，则称 (X,Y) 服从参数为 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的**二维正态分布**. 记为 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

二维正态分布的概率密度 $f(x,y)$ 在三维空间的图形好像是一个椭圆切面的钟倒扣在 xOy 平面上，其中心在 (μ_1,μ_2) 处，如图 3-4 所示.

3.2　边缘分布

有了二维随机变量 (X,Y) 的联合分布，有时仍然需要了解 X,Y 各自的分布，以及 (X,Y) 的联合分布与 X,Y 各自分布之间的关系. 本节讨论这个问题.

3.2.1　边缘分布函数

定义 3.6　设 $F(x,y)$ 为二维随机变量 (X,Y) 的联合分布函数，则随机变量 X 的分布函数

$$F_X(x)=P\{X\leqslant x\}=P\{X\leqslant x,Y<+\infty\}=F(x,+\infty)$$

称为二维随机变量(X,Y)关于随机变量 X 的**边缘分布函数**.

同理,随机变量 Y 的分布函数 $F_Y(y)=F(+\infty,y)$,称为二维随机变量(X,Y)关于 Y 的**边缘分布函数**.

3.2.2 离散型随机变量的边缘分布律

定义 3.7 设二维离散型随机变量(X,Y)的联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

则(X,Y)关于随机变量 X 的分布律为

$$P\{X=x_i\}=P\{X=x_i,Y<+\infty\}=\sum_{j=1}^{+\infty}p_{ij},\quad i=1,2,3,\cdots.$$

通常,称其为(X,Y)关于 X 的**边缘分布律**,记为 $p_{i\cdot}$,即

$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{\infty}p_{ij},\quad i=1,2,3,\cdots. \tag{3.7}$$

同理,(X,Y)关于 Y 的**边缘分布律**为

$$p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{\infty}p_{ij},\quad j=1,2,3,\cdots. \tag{3.8}$$

人们常常习惯于把二维离散型随机变量(X,Y)的联合分布律和边缘分布律用表格来表示,如下所示.

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	

将边缘分布律记在联合分布律表的边上,既容易计算又一目了然,这也是“边缘”二字的含义.

例 1 (续 3.1 节例 1)试求(X,Y)的边缘分布律.

解 见下表

X \ Y	1	2	3	4	$p_{i\cdot}$
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$	1

例 2 袋中装有2只白球和3只黑球，现连续摸球两次，定义下列随机变量：

$$X=\begin{cases}1 & 第一次摸出白球\\0 & 第一次摸出黑球\end{cases},\quad Y=\begin{cases}1 & 第二次摸出白球\\0 & 第二次摸出黑球\end{cases}.$$

分别就有放回摸球与无放回摸球两种方式，求(X,Y)的联合分布律和边缘分布律.

解 (1) 有放回摸球情形

由试验的独立性可得

$$P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\}=\frac{3}{5}\times\frac{3}{5}.$$

类似可求出其他的 p_{ij}，最后得(X,Y)的联合分布律和边缘分布律，见下表：

X \ Y	0	1	$p_{i\cdot}$
0	$\frac{3}{5}\times\frac{3}{5}$	$\frac{3}{5}\times\frac{2}{5}$	$\frac{3}{5}$
1	$\frac{2}{5}\times\frac{3}{5}$	$\frac{2}{5}\times\frac{2}{5}$	$\frac{2}{5}$
$p_{\cdot j}$	$\frac{3}{5}$	$\frac{2}{5}$	1

(2) 无放回摸球情形

由概率乘法公式可得

$$P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\mid X=0\}=\frac{3}{5}\times\frac{2}{4}.$$

类似可求出其他的 p_{ij}，最后得(X,Y)的联合分布律和边缘分布律，见下表：

X \ Y	0	1	$p_{i\cdot}$
0	$\frac{3}{5}\times\frac{2}{4}$	$\frac{3}{5}\times\frac{2}{4}$	$\frac{3}{5}$
1	$\frac{2}{5}\times\frac{3}{4}$	$\frac{2}{5}\times\frac{1}{4}$	$\frac{2}{5}$
$p_{\cdot j}$	$\frac{3}{5}$	$\frac{2}{5}$	1

上面两种情形中 X 和 Y 的边缘分布律是相同的，但它们的联合分布律却完全不同，由此可知，仅由边缘分布律未必能确定联合分布律.

3.2.3 连续型随机变量的边缘概率密度

定义 3.8 对于连续型随机变量(X,Y)，设它的联合概率密度为 $f(x,y)$，由

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(u,v)\mathrm{d}v\right]\mathrm{d}u$$

知 X 是一个连续型随机变量，且其概率密度为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,\tag{3.9}$$

称为随机变量(X,Y)关于 X 的**边缘概率密度**.

同理，将

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x \tag{3.10}$$

称为随机变量(X,Y)关于Y的**边缘概率密度**.

例 3 设二维随机变量(X,Y)具有概率密度

$$f(x,y)=\begin{cases}48xy, & 0<x<1,x^3<y<x^2\\ 0, & \text{其他}\end{cases},$$

求(X,Y)的边缘概率密度.

解

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{x^3}^{x^2}48xy\mathrm{d}y, & 0<x<1\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}24(x^5-x^7), & 0<x<1\\ 0, & \text{其他}\end{cases};$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_{\sqrt{y}}^{\sqrt[3]{y}}48xy\mathrm{d}x, & 0<y<1\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}24(y^{\frac{5}{3}}-y^2), & 0<y<1\\ 0, & \text{其他}\end{cases}.$$

例 4 设二维随机变量(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{4y}{\pi(1+x^2)}, & x>0,0\leqslant y\leqslant 1\\ 0, & \text{其他}\end{cases},$$

求(X,Y)的边缘概率密度函数$f_X(x)$,$f_Y(y)$.

解

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_0^1\dfrac{4y}{\pi(1+x^2)}\mathrm{d}y, & x>0,\\ 0, & x\leqslant 0\end{cases}$$

$$=\begin{cases}\dfrac{2}{\pi(1+x^2)}, & x>0\\ 0, & x\leqslant 0\end{cases};$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_0^{+\infty}\dfrac{4y}{\pi(1+x^2)}\mathrm{d}x, & 0\leqslant y\leqslant 1,\\ 0, & \text{其他}\end{cases}$$

$$=\begin{cases}2y, & 0\leqslant y\leqslant 1\\ 0, & \text{其他}\end{cases}.$$

例 5 设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,求$(X,Y)$的边缘概率密度.

解

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]},$$

令$\frac{x-\mu_1}{\sigma_1}=u$，$\frac{y-\mu_2}{\sigma_2}=v$，则

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}\mathrm{d}y \\
&= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)}\mathrm{d}v \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{u^2}{2}}\frac{1}{\sqrt{2\pi(1-\rho^2)}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{\rho^2u^2-2\rho uv+v^2}{2(1-\rho^2)}}\mathrm{d}v \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{u^2}{2}}\frac{1}{\sqrt{2\pi(1-\rho^2)}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{(v-\rho u)^2}{2(1-\rho^2)}}\mathrm{d}v \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{u^2}{2}} = \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}.
\end{aligned}
$$

可知 $X\sim N(\mu_1,\sigma_1^2)$，由对称性可得

$$
f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}.
$$

即 $Y\sim N(\mu_2,\sigma_2^2)$. 因而正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的两个边缘分布均为一维正态分布，分别为 $N(\mu_1,\sigma_1^2)$和 $N(\mu_2,\sigma_2^2)$，它们与参数 ρ 无关.

*3.3 二维随机变量的条件分布

在对二维随机变量(X,Y)的讨论中，有时需考虑其中一个随机变量取定某个值时，另一个随机变量的概率分布. 借助于随机事件的条件概率的概念，我们引入随机变量的条件分布.

3.3.1 离散型随机变量的条件分布

设二维离散型随机变量(X,Y)的分布律为

$$
P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots.
$$

仿照条件概率的定义，我们可以很容易地给出离散型随机变量的条件分布律.

定义 3.9 设(X,Y)是二维离散型随机变量，对于固定的 j，若 $P\{Y=y_j\}>0$，则称

$$
p_{i|j}=P\{X=x_i\mid Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\cdots \tag{3.11}
$$

为在 $Y=y_j$ 条件下随机变量 X 的**条件分布律**.

同样，对于固定的 i，若 $P\{X=x_i\}>0$，则称

$$
p_{j|i}=P\{Y=y_j\mid X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}},\quad j=1,2,\cdots \tag{3.12}
$$

为在 $X=x_i$ 条件下随机变量 Y 的**条件分布律**.

有了条件分布律，就可以给出离散型随机变量的条件分布函数.

定义 3.10 给定 $Y=y_j$ 条件下 X 的**条件分布函数**为

$$F(x \mid y_j) = P\{X \leqslant x \mid Y = y_j\} = \sum_{x_i \leqslant x} P\{X = x_i \mid Y = y_j\} = \sum_{x_i \leqslant x} p_{i|j}, \quad (3.13)$$

给定 $X=x_i$ 条件下 Y 的**条件分布函数**为

$$F(y \mid x_i) = P\{Y \leqslant y \mid X = x_i\} = \sum_{y_j \leqslant y} P\{Y = y_j \mid X = x_i\} = \sum_{y_j \leqslant y} p_{j|i} \quad (3.14)$$

例 1 在 3.2 节例 2 的无放回摸球情形中，分别求在条件 $X=0$ 和 $X=1$ 下随机变量 Y 的条件分布.

解 在 3.2 节例 2 中，已求得 (X,Y) 的联合分布律及边缘分布律为

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{5}$
1	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{2}{5}$
$p_{\cdot j}$	$\frac{3}{5}$	$\frac{2}{5}$	1

又由

$$P\{Y = 0 \mid X = 0\} = \frac{P\{X = 0, Y = 0\}}{P\{X = 0\}} = \frac{\frac{3}{10}}{\frac{3}{5}} = \frac{1}{2},$$

$$P\{Y = 1 \mid X = 0\} = \frac{P\{X = 0, Y = 1\}}{P\{X = 0\}} = \frac{\frac{3}{10}}{\frac{3}{5}} = \frac{1}{2},$$

可得在条件 $X=0$ 下随机变量 Y 的条件分布律为

$Y \mid X=0$	0	1
p	$\frac{1}{2}$	$\frac{1}{2}$

同理，在条件 $X=1$ 下随机变量 Y 的条件分布律为

$Y \mid X=1$	0	1
p	$\frac{3}{4}$	$\frac{1}{4}$

从本例题看出，每个条件分布都从一个侧面描述了一种状态下的特定分布.

3.3.2 连续型随机变量的条件分布

设二维连续型随机变量 (X,Y) 的联合密度函数为 $f(x,y)$，边缘概率密度为 $f_X(x)$，$f_Y(y)$.

在离散型随机变量场合，条件分布函数为 $P\{X\leqslant x|Y=y\}$，但是，因为连续型随机变量取某个值的概率为零，即 $P\{Y=y\}=0$，所以无法用条件概率直接计算 $P\{X\leqslant x|Y=y\}$. 这

时，我们可以将 $P\{X\leqslant x|Y=y\}$看成是 $h\to 0$ 时 $P\{X\leqslant x|y\leqslant Y\leqslant y+h\}$的极限，即

$$\begin{aligned}P\{X\leqslant x\mid Y=y\}&=\lim_{h\to 0}P\{X\leqslant x\mid y\leqslant Y\leqslant y+h\}\\&=\lim_{h\to 0}\frac{P\{X\leqslant x,y\leqslant Y\leqslant y+h\}}{P\{y\leqslant Y\leqslant y+h\}}\\&=\lim_{h\to 0}\frac{\int_{-\infty}^{x}\int_{y}^{y+h}f(u,v)\mathrm{d}v\mathrm{d}u}{\int_{y}^{y+h}f_Y(v)\mathrm{d}v}\\&=\lim_{h\to 0}\frac{\int_{-\infty}^{x}\left[\frac{1}{h}\int_{y}^{y+h}f(u,v)\mathrm{d}v\right]\mathrm{d}u}{\frac{1}{h}\int_{y}^{y+h}f_Y(v)\mathrm{d}v}.\end{aligned}$$

当 $f_Y(y)$，$f(x,y)$在 y 处连续时，由积分中值定理可得

$$\lim_{h\to 0}\frac{1}{h}\int_{y}^{y+h}f_Y(v)\mathrm{d}v=f_Y(y),$$

$$\lim_{h\to 0}\frac{1}{h}\int_{y}^{y+h}f(u,v)\mathrm{d}v=f(u,y).$$

所以

$$P\{X\leqslant x\mid Y=y\}=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}\mathrm{d}u.$$

至此，我们可以定义连续型随机变量的条件分布如下.

定义 3.11 设(X,Y)是二维连续型随机变量，对固定的 y，若 $f_Y(y)>0$，则称

$$F(x\mid y)=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}\mathrm{d}u \tag{3.15}$$

为在 $Y=y$ 条件下 X 的**条件分布函数**，

$$f(x\mid y)=\frac{f(x,y)}{f_Y(y)} \tag{3.16}$$

为在 $Y=y$ 条件下 X 的**条件概率密度**.

同理，对固定的 x，若 $f_X(x)>0$，则称

$$F(y\mid x)=\int_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}\mathrm{d}v \tag{3.17}$$

为在 $X=x$ 条件下 Y 的**条件分布函数**，

$$f(y\mid x)=\frac{f(x,y)}{f_X(x)} \tag{3.18}$$

为在 $X=x$ 条件下 Y 的**条件概率密度**.

例 2 设(X,Y)服从 $G=\{(x,y)|x^2+y^2\leqslant 1\}$上的均匀分布，试求给定 $Y=y$ 条件下的条件概率密度 $f(x|y)$.

解 因为

$$f(x,y)=\begin{cases}\frac{1}{\pi} & x^2+y^2\leqslant 1\\0, & \text{其他}\end{cases},$$

由此得 Y 的边缘概率密度为

$$f_Y(y)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2}, & -1\leqslant y\leqslant 1\\ 0, & \text{其他}\end{cases}.$$

所以当$-1<y<1$时,有

$$f(x\mid y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{\dfrac{1}{\pi}}{\dfrac{2}{\pi}\sqrt{1-y^2}}=\dfrac{1}{2\sqrt{1-y^2}}, & -\sqrt{1-y^2}\leqslant x\leqslant\sqrt{1-y^2}\\ 0, & \text{其他}\end{cases}.$$

将$y=0$和$y=0.5$分别代入上式可得(两个均匀分布)

$$f(x\mid y=0)=\begin{cases}\dfrac{1}{2}, & -1\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases},$$

$$f(x\mid y=0.5)=\begin{cases}\dfrac{1}{\sqrt{3}}, & -\dfrac{\sqrt{3}}{2}\leqslant x\leqslant\dfrac{\sqrt{3}}{2}\\ 0, & \text{其他}\end{cases}.$$

进一步有:当$-1<y<1$时给定$Y=y$条件下,X服从$(-\sqrt{1-y^2},\sqrt{1-y^2})$上的均匀分布;同理,当$-1<x<1$时给定$X=x$条件下,$Y$服从$(-\sqrt{1-x^2},\sqrt{1-x^2})$上的均匀分布.

3.4 随机变量的独立性

随机变量的独立性是概率论中的一个重要概念.借助于随机事件的独立性,我们引入随机变量独立性的概念.

定义 3.12 若二维随机变量(X,Y)对于任意实数x,y都有

$$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\},\tag{3.19}$$

则称随机变量X与Y是**相互独立**的.

若二维随机变量(X,Y)的分布函数为$F(x,y)$,边缘分布函数为$F_X(x)$和$F_Y(y)$,则由式(3.19)得随机变量X与Y相互独立等价于对任意实数x,y都有$F(x,y)=F_X(x)F_Y(y)$.

当(X,Y)是离散型随机变量时,X与Y相互独立的条件等价于对(X,Y)的所有可能取的数偶(x_i,y_j),有

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\},\quad \text{即}\quad p_{ij}=p_{i\cdot}p_{\cdot j}.$$

当(X,Y)是连续型随机变量时,X与Y相互独立的条件等价于

$$f(x,y)=f_X(x)f_Y(y),$$

其中$f(x,y)$和$f_X(x),f_Y(y)$分别为(X,Y)的联合概率密度和边缘概率密度.

在随机变量X与Y相互独立的情况下,(X,Y)的联合分布与边缘分布可以相互确定.

可以证明,若X与Y相互独立,则$f(X)$和$g(Y)$也相互独立.

例 1 在3.2节例2中,有放回摸球情形,由于对任意的i,j有$p_{ij}=p_{i\cdot}p_{\cdot j}$,所以$X$与$Y$是相互独立的.无放回摸球情形,由于

$$P\{X=0,Y=0\}=\frac{3}{5}\times\frac{4}{5}\neq\frac{3}{5}\times\frac{3}{5}=P\{X=0\}P\{Y=0\},$$

所以 X 与 Y 不是相互独立的.

例 2 设两个独立的随机变量 X 与 Y 的分布律为

X	1	3
p_X	0.3	0.7

Y	2	4
p_Y	0.6	0.4

求随机变量(X,Y)的联合分布律.

解 因为 X 与 Y 是相互独立的,所以对于任意的 i,j,有

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}.$$

于是,

$$P\{X=1,Y=2\}=P\{X=1\}P\{Y=2\}=0.3\times0.6=0.18,$$
$$P\{X=1,Y=4\}=P\{X=1\}P\{Y=4\}=0.3\times0.4=0.12,$$
$$P\{X=3,Y=2\}=P\{X=3\}P\{Y=2\}=0.7\times0.6=0.42,$$
$$P\{X=3,Y=4\}=P\{X=3\}P\{Y=4\}=0.7\times0.4=0.28.$$

因此(X,Y)的联合分布律为

X \ Y	2	4
1	0.18	0.12
3	0.42	0.28

例 3 在3.2节例3中判断 X,Y 是否相互独立.

解 由3.2节例3知

$$f_X(x)=\begin{cases}24(x^5-x^7), & 0<x<1\\ 0 & 其他\end{cases},$$

$$f_Y(y)=\begin{cases}24(y^{\frac{5}{3}}-y^2), & 0<y<1\\ 0 & 其他\end{cases},$$

因为 $f_X(x)f_Y(y)\neq f(x,y)$,所以 X 与 Y 不相互独立.

例 4 设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,试证明 X 与 Y 相互独立的充要条件是 $\rho=0$.

证明 充分性:若 $\rho=0$,由3.2节例4知,对任意实数 x,y,有

$$f(x,y)=f_X(x)f_Y(y),$$

所以 X 与 Y 相互独立.

必要性:设 X 与 Y 相互独立,则对任意的实数 x,y,都有

$$f(x,y)=f_X(x)f_Y(y),$$

即

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$

$$=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]},\quad -\infty<x<+\infty,-\infty<y<+\infty.$$

显然,当 $x=\mu_1,y=\mu_2$ 时上式仍成立,代入得

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\frac{1}{2\pi\sigma_1\sigma_2},$$

由此得 $\rho=0$.

以上关于二维随机变量的一些概念很容易推广到 n 维随机变量的情况.

1) 分布函数

n 维随机变量$(X_1,X_2,\cdots,X_n)$的分布函数

$$F(x_1,x_2,\cdots,x_n)=P\{X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n\},$$

其中 $x_1,x_2,\cdots,x_n$ 为任意实数.

2) 概率密度函数

若存在非负函数 $f(x_1,x_2,\cdots,x_n)$,使对于任意实数 $x_1,x_2,\cdots,x_n$ 有

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}}\cdots\int_{-\infty}^{x_1}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n,$$

则称 $f(x_1,x_2,\cdots,x_n)$为$(X_1,X_2,\cdots,X_n)$的概率密度函数.

3) 边缘分布函数

$$F_{X_1}(x_1)=F(x_1,+\infty,+\infty,\cdots,+\infty)$$

称为 n 维随机变量$(X_1,X_2,\cdots,X_n)$关于 X_1 的边缘分布函数,

$$F_{X_1,X_2}(x_1,x_2)=F(x_1,x_2,+\infty,\cdots,+\infty)$$

称为 n 维随机变量$(X_1,X_2,\cdots,X_n)$关于(X_1,X_2)的边缘分布函数.

其他以此类推.

4) 边缘概率密度函数

若 $f(x_1,x_2,\cdots,x_n)$是$(X_1,X_2,\cdots,X_n)$的概率密度,则$(X_1,X_2,\cdots,X_n)$关于 X_1、关于(X_1,X_2)的边缘概率密度分别为

$$f_{X_1}(x_1)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_2\mathrm{d}x_3\cdots\mathrm{d}x_n,$$

$$f_{X_1,X_2}(x_1,x_2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_3\mathrm{d}x_4\cdots\mathrm{d}x_n.$$

同理可得$(X_1,X_2,\cdots,X_n)$的 $k(1\leqslant k<n)$维边缘概率密度.

5) 相互独立性

若对于所有的 $x_1,x_2,\cdots,x_n$ 有

$$F(x_1,x_2,\cdots,x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n),$$

则称 $X_1,X_2,\cdots,X_n$ 是相互独立的.

若对于所有的 $x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_n$ 有

$$F(x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_n)=F_1(x_1,x_2,\cdots,x_m)F_2(y_1,y_2,\cdots,y_n),$$

其中 F_1,F_2,F 依次为随机变量 $(X_1,X_2,\cdots,X_m)$,$(Y_1,Y_2,\cdots,Y_n)$和$(X_1,X_2,\cdots,X_m,Y_1,Y_2,\cdots,Y_n)$的分布函数,则称随机变量$(X_1,X_2,\cdots,X_m)$与$(Y_1,Y_2,\cdots,Y_n)$相互独立.

进一步有：设$(X_1,X_2,\cdots,X_m)$和$(Y_1,Y_2,\cdots,Y_n)$相互独立，则$X_i(i=1,2,\cdots,m)$和$Y_j(j=1,2,\cdots,n)$相互独立. 又若h,g是连续函数，则$h(X_1,X_2,\cdots,X_m)$和$g(Y_1,Y_2,\cdots,Y_n)$相互独立.

3.5 二维随机变量的函数的分布

在第2章中我们已经讨论过一维随机变量的函数的分布，本节讨论二维随机变量的函数的分布.

3.5.1 二维离散型随机变量的函数的分布

设二维离散型随机变量(X,Y)的联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

则函数$Z=g(X,Y)$是一维离散型随机变量. 设$Z=g(X,Y)$的所有不同的取值为$z_1,z_2,\cdots,z_k,\cdots$，则$Z=g(X,Y)$的分布律为

$$P\{Z=z_k\}=\sum_{g(x_i,y_j)=z_k} p_{ij},\quad k=1,2,\cdots. \tag{3.20}$$

例1 设(X,Y)的联合分布律为

X \ Y	-1	1	2
-1	$\frac{5}{20}$	$\frac{2}{20}$	$\frac{6}{20}$
2	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{1}{20}$

求：(1)$Z_1=X+Y$的分布律；(2)$Z_2=\max\{X,Y\}$的分布律.

解 (1) 将(X,Y)及各个函数的取值对应列于同一表中：

p	$\frac{5}{20}$	$\frac{2}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{1}{20}$
(X,Y)	$(-1,-1)$	$(-1,1)$	$(-1,2)$	$(2,-1)$	$(2,1)$	$(2,2)$
$Z_1=X+Y$	-2	0	1	1	3	4
$Z_2=\max\{X,Y\}$	-1	1	2	2	2	2

合并整理后可得

$Z_1=X+Y$	-2	0	1	3	4
p	$\frac{5}{20}$	$\frac{2}{20}$	$\frac{9}{20}$	$\frac{3}{20}$	$\frac{1}{20}$

$Z_2=\max\{X,Y\}$	-1	1	2
p	$\frac{5}{20}$	$\frac{2}{20}$	$\frac{13}{20}$

3.5.2 二维连续型随机变量的函数的分布

1. $Z=X+Y$的分布

设二维随机变量(X,Y)的密度函数为 $f(x,y)$，则 $Z=X+Y$ 的分布函数

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}=\iint_{x+y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\mathrm{d}y\int_{-\infty}^{z-y}f(x,y)\mathrm{d}x=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z-y}f(x,y)\mathrm{d}x\right)\mathrm{d}y,$$

由此可得 Z 的概率密度为

$$f_Z(z)=\frac{\mathrm{d}F_Z(z)}{\mathrm{d}z}=\int_{-\infty}^{+\infty}f(z-y,y)\mathrm{d}y. \tag{3.21}$$

利用 X 和 Y 的对称性，可得

$$f_Z(z)=\frac{\mathrm{d}F_Z(z)}{\mathrm{d}z}=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x. \tag{3.22}$$

特别地，当 X 与 Y 相互独立时，有

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\mathrm{d}y=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x, \tag{3.23}$$

其中 $f_X(x)$和 $f_Y(y)$分别为 X 和 Y 的边缘概率密度.

式(3.23)又称为**卷积公式**，并记为 $f_Z(z)=f_X*f_Y$.

例 2 设二维随机变量(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}24y(1-x), & 0\leqslant x\leqslant 1,0\leqslant y\leqslant x\\ 0, & \text{其他}\end{cases},$$

求 $Z=X+Y$ 的概率密度函数 $f_Z(z)$.

解 $f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x$，由 $f(x,y)$ 知 $f(x,z-x)\neq 0$ 范围为$\begin{cases}0\leqslant x\leqslant 1\\ 0\leqslant z-x\leqslant x\end{cases}$，见图 3-5.

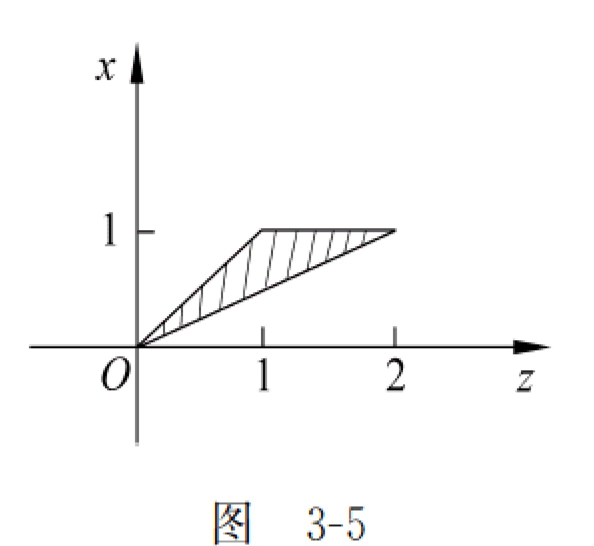


图 3-5

于是

$$f_Z(z)=\begin{cases}0, & z<0\\ \int_{\frac{z}{2}}^{z}24(z-x)(1-x)\mathrm{d}x, & 0\leqslant z<1\\ \int_{\frac{z}{2}}^{1}24(z-x)(1-x)\mathrm{d}x, & 1\leqslant z<2\\ 0, & z\geqslant 2\end{cases}$$

$$=\begin{cases}3z^2-2z^3, & 0\leqslant z<1\\ 2z^3-9z^2+12z-4, & 1\leqslant z<2.\\ 0, & \text{其他}\end{cases}$$

例 3 设 $X\sim N(0,1)$，$Y\sim N(0,1)$，且 X 与 Y 相互独立，求 $Z=X+Y$ 的概率密度.

解 由卷积公式,得

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{e}^{-\frac{(z-x)^2}{2}}\mathrm{d}x$$
$$=\frac{1}{2\pi}\mathrm{e}^{-\frac{z^2}{4}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\left(x-\frac{z}{2}\right)^2}\mathrm{d}x.$$

令 $t=x-\frac{z}{2}$ 得

$$f_Z(z)=\frac{1}{2\pi}\mathrm{e}^{-\frac{z^2}{4}}\int_{-\infty}^{+\infty}\mathrm{e}^{-t^2}\mathrm{d}t=\frac{1}{2\sqrt{\pi}}\mathrm{e}^{-\frac{z^2}{4}}=\frac{1}{\sqrt{2\pi}\times\sqrt{2}}\mathrm{e}^{-\frac{(z-0)^2}{2\times(\sqrt{2}^2)}},$$

即 $Z\sim N(0,2)$.

一般地,设 X,Y 相互独立,且 $X\sim N(\mu_1,\sigma_1^2)$,$Y\sim N(\mu_2,\sigma_2^2)$,则 $Z=X+Y$ 仍然服从正态分布,且有 $Z\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

更一般地,设 $X_1,X_2,\cdots,X_n$ 相互独立,且 $X_i\sim N(\mu_i,\sigma_i^2)$,$i=1,2,\cdots,n$,则对于任意不全为零的常数 $c_1,c_2,\cdots,c_n$,有

$$\sum_{i=1}^{n}c_iX_i\sim N\left(\sum_{i=1}^{n}c_i\mu_i,\sum_{i=1}^{n}c_i^2\sigma_i^2\right).$$

2. $Z_1=\max\{X,Y\}$, $Z_2=\min\{X,Y\}$ 的分布

设 X,Y 是两个相互独立的随机变量,它们的分布函数分别为 $F_X(x)$,$F_Y(y)$. 现在来求 Z_1 与 Z_2 的分布函数.

对于任意实数 z,由于 $\{Z_1\leqslant z\}=\{X\leqslant z,Y\leqslant z\}$,且 X,Y 独立,因此

$$F_{\max}(z)=P\{Z_1\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}=P\{X\leqslant z\}P\{Y\leqslant z\},$$

即有

$$F_{\max}(z)=F_X(z)F_Y(z). \tag{3.24}$$

类似地,可得 Z_2 的分布函数为

$$F_{\min}(z)=P\{Z_2\leqslant z\}=1-P\{Z_2>z\}$$
$$=1-P\{X>z,Y>z\}=1-P\{X>z\}P\{Y>z\}$$
$$=1-[1-P\{X\leqslant z\}][1-P\{Y\leqslant z\}]$$
$$=1-[1-F_X(z)][1-F_Y(z)],$$

即

$$F_{\min}(z)=1-[1-F_X(z)][1-F_Y(z)]. \tag{3.25}$$

以上结果可推广到任意 n 个相互独立的随机变量的情况.

设 $X_1,X_2,\cdots,X_n$ 是 n 个相互独立的随机变量,它们的分布函数分别为 $F_{X_1}(x_1)$,$F_{X_2}(x_2)$,$\cdots$,$F_{X_n}(x_n)$,则 $Z_1=\max\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_{\max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z), \tag{3.26}$$

$Z_2=\min\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_{\min}(z)=1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]. \tag{3.27}$$

特别地,当 $X_1,X_2,\cdots,X_n$ 相互独立且具有相同的分布函数 $F(x)$ 时,有

$$F_{\max}(z)=[F(z)]^n, \tag{3.28}$$

$$F_{\min}(z)=1-[1-F(z)]^n. \tag{3.29}$$

例 4 设系统 L 由两个独立工作的电子原件 L_1,L_2 连接而成. 连接的方式分别为：(1)串联；(2)并联(见图 3-6). 设 L_1,L_2 的寿命分别为 X,Y，已知它们的概率密度分别为

$$f_X(x)=\begin{cases}\theta_1 e^{-\theta_1 x}, & x>0\\ 0, & x\leqslant 0\end{cases},\quad f_Y(y)=\begin{cases}\theta_2 e^{-\theta_2 y}, & y>0\\ 0, & y\leqslant 0\end{cases},$$

其中 $\theta_1>0,\theta_2>0$. 试分别就以上两种连接方式求系统的寿命 Z 的概率密度.

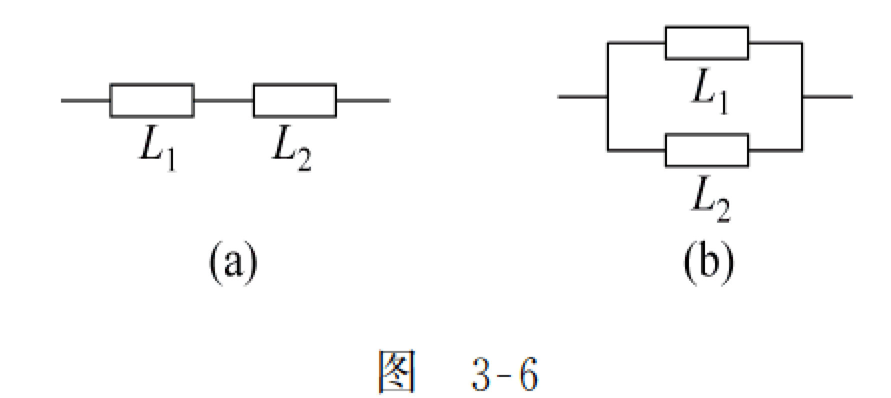


图 3-6

解 (1) 串联的情况

由于当 L_1,L_2 中有一个损坏时，系统 L 就停止工作，因而系统 L 的寿命

$$Z=\min\{X,Y\}.$$

由题设，X,Y 的分布函数分别为

$$F_X(x)=\begin{cases}1-e^{-\theta_1 x}, & x>0\\ 0, & x\leqslant 0\end{cases},\quad F_Y(y)=\begin{cases}1-e^{-\theta_2 y}, & y>0\\ 0, & y\leqslant 0\end{cases},$$

于是 $Z=\min\{X,Y\}$ 的分布函数为

$$F_{\min}(z)=1-[1-F_X(z)][1-F_Y(z)]=\begin{cases}1-e^{-(\theta_1+\theta_2)z}, & z>0\\ 0, & z\leqslant 0\end{cases};$$

$Z=\min\{X,Y\}$ 的概率密度为

$$f_{\min}(z)=\begin{cases}(\theta_1+\theta_2)e^{-(\theta_1+\theta_2)z}, & z>0\\ 0, & z\leqslant 0\end{cases}.$$

(2) 并联的情况

由于当 L_1,L_2 都损坏时系统 L 才停止工作，所以 L 的寿命 Z 为

$$Z=\max\{X,Y\}.$$

于是 $Z=\max\{X,Y\}$ 的分布函数为

$$F_{\max}(z)=F_X(z)F_Y(z)=\begin{cases}(1-e^{-\theta_1 z})(1-e^{-\theta_2 z}), & z>0\\ 0, & z\leqslant 0\end{cases};$$

$Z=\max\{X,Y\}$ 的概率密度为

$$f_{\max}(z)=\begin{cases}\theta_1 e^{-\theta_1 z}+\theta_2 e^{-\theta_2 z}-(\theta_1+\theta_2)e^{-(\theta_1+\theta_2)z}, & z>0\\ 0, & z\leqslant 0\end{cases}.$$

小　　结

本章主要介绍了二维随机变量及其概率分布，主要内容有：二维随机变量及其联合分布函数，二维离散型随机变量的联合分布律，边缘分布律，条件概率分布，二维连续型随机变量的联合概率密度函数，边缘概率密度函数，条件概率密度函数，随机变量的独立性，两个随机变量之和的分布，两个独立随机变量最大值、最小值的分布.

基本要求：

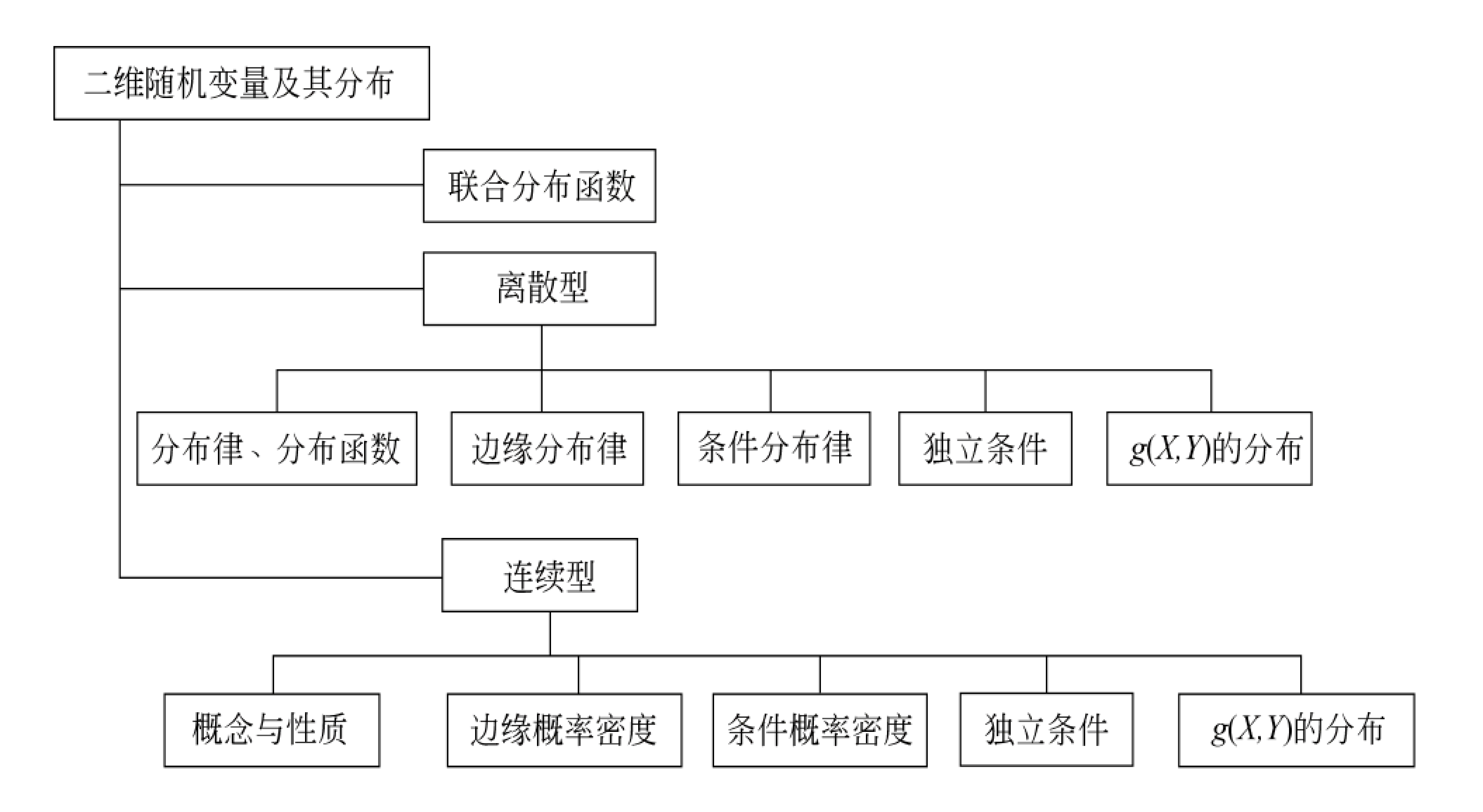
(1) 熟练掌握二维随机变量及其联合分布函数，二维离散型随机变量的联合分布律，边缘分布律，二维连续型随机变量的联合概率密度函数，边缘概率密度函数，随机变量的独立性等概念.

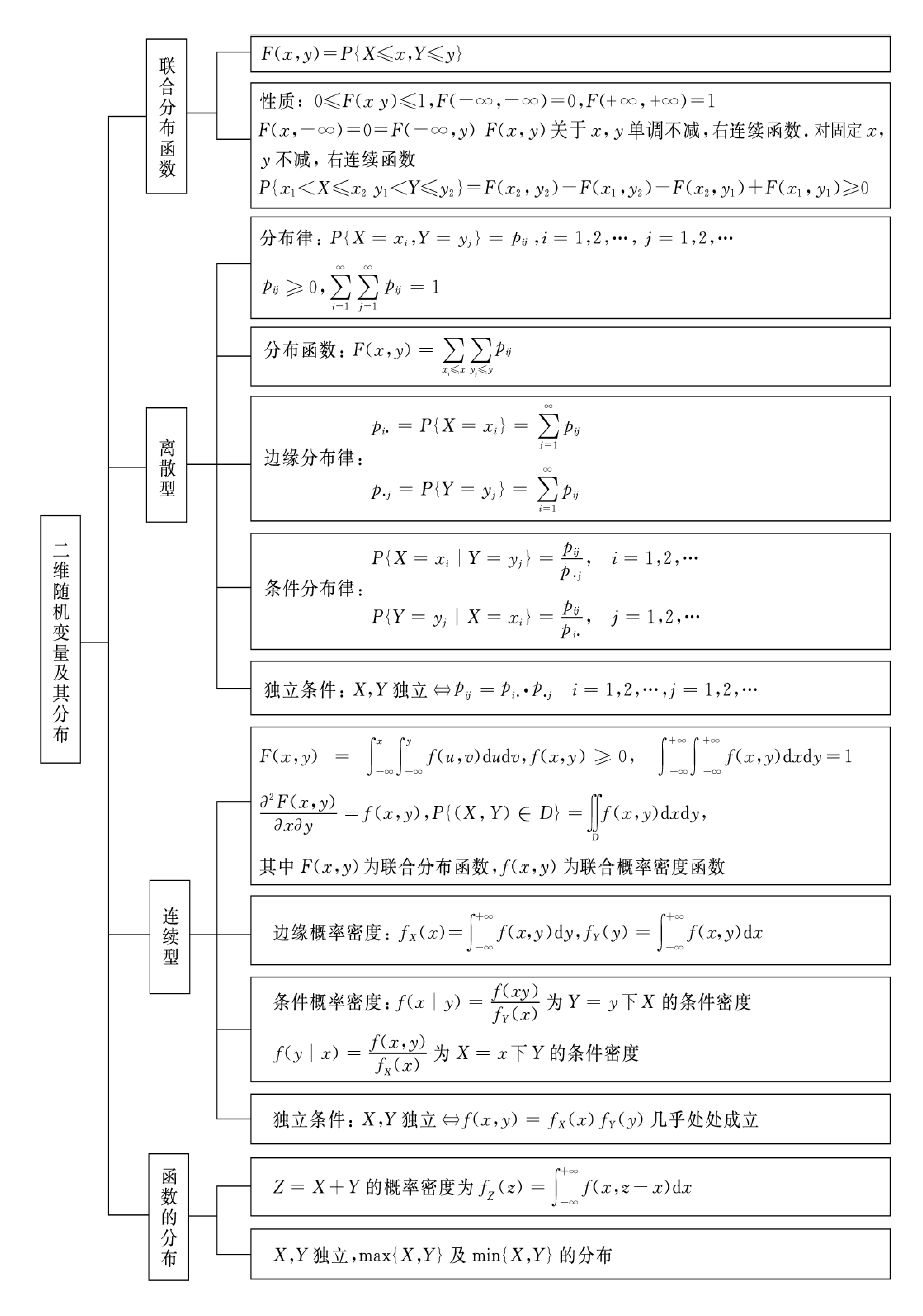
(2) 会利用联合分布函数求联合概率密度函数，利用联合概率密度函数求联合分布函数，会求边缘分布函数，熟练掌握求边缘概率密度函数或边缘分布律的方法，会求二维连续型随机变量之和的概率密度函数，会求两个独立随机变量最大值、最小值的分布.

重点： 二维随机变量的概念，联合分布函数，联合分布律，联合概率密度函数，边缘分布函数，边缘分布律，边缘概率密度函数，随机变量的独立性.

难点： 已知联合概率密度函数求联合分布函数，边缘概率密度函数，$Z=X+Y$ 的概率密度函数.

本章知识结构：





习 题 3

1. 某盒中有形状相同的 a 个白球和 b 个黑球，每次从中任选一球，共取两次，设 X 及 Y 分别表示第一次及第二次取出的黑球数. 分别就(1)放回抽样(2)不放回抽样两种情况，求二维随机变量(X,Y)的联合分布.

2. 将一枚均匀的硬币抛掷 3 次，设 X 为 3 次抛掷中正面出现的次数，Y 为反面出现的次数，求 (X,Y) 的联合分布律.

3. 设随机变量 (X,Y) 的联合分布函数为

$$F(x,y)=(A+B\arctan x)(A+B\arctan y)\left[1+\frac{1}{2}(A-B\arctan x)(A-B\arctan y)\right].$$

求：(1)常数 A,B；(2) $P\{X\leqslant 0,Y\leqslant 0\}$.

4. 设随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}k\mathrm{e}^{-(3x+4y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases}.$$

求：(1) 常数 k；

(2) (X,Y) 的联合分布函数 $F(x,y)$；

(3) $P\left\{X<\frac{1}{3},Y>\frac{1}{4}\right\}$.

5. 已知随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}kx^2y, & 0<x<y<1\\ 0, & \text{其他}\end{cases}.$$

求：(1)常数 k；(2) $P\{X+Y\leqslant 1\}$.

6. 设某仪器由两个部件构成，X 与 Y 分别是这两个部件的寿命(千小时). 已知 (X,Y) 的联合分布函数为

$$F(x,y)=\begin{cases}1-\mathrm{e}^{-0.5x}-\mathrm{e}^{-0.5y}+\mathrm{e}^{-0.5(x+y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases}.$$

求：(1) 边缘分布函数 $F_X(x),F_Y(y)$；

(2) 联合概率密度 $f(x,y)$ 及边缘密度 $f_X(x),f_Y(y)$；

(3) 两部件寿命均超过 100 h 的概率.

7. 设二维离散随机变量 (X,Y) 的可能取值为 $(0,0),(-1,1),(-1,2),(1,0)$，且取这些值的概率依次为 $\frac{1}{6},\frac{1}{3},\frac{1}{12},\frac{5}{12}$. 试求 X 与 Y 各自的边缘分布律.

8. 已知 (X,Y) 的联合分布律为

X \ Y	1	2	3
1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
2	$\frac{2}{12}$	$\frac{1}{12}$	0
3	$\frac{2}{12}$	$\frac{1}{12}$	0
4	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

求：(1) X 和 Y 的边缘分布律；

(2) 在 $X=4$ 下 Y 的条件分布律和 $Y=3$ 下 X 的分布律.

9. 设随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}1, & |y|<x,0<x<1\\0, & \text{其他}\end{cases},$$

求条件概率密度 $f(x|y)$.

10. 已知随机变量 Y 的概率密度为 $f_Y(y)=\begin{cases}5y^4, & 0<y<1\\0, & \text{其他}\end{cases}$,又设 X 关于 Y 的条件概率密度为

$$f(x\mid y)=\begin{cases}\dfrac{3x^2}{y^3}, & 0<x<y\\0, & \text{其他}\end{cases},$$

求 $P\{X>0.5\}$.

11. 已知随机变量 X 和 Y 的概率分布律为

X	-1	0	1
p_X	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Y	0	1
p_Y	$\frac{1}{2}$	$\frac{1}{2}$

且 $P\{XY=0\}=1$.

(1) 求 (X,Y) 的联合分布律;

(2) X 与 Y 是否独立? 并说明理由.

12. 已知随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{3}{2}x^2\mathrm{e}^{-y}, & -1<x<1,y>0\\0, & \text{其他}\end{cases}.$$

(1) 求边缘概率密度 $f_X(x),f_Y(y)$;

(2) X 与 Y 是否相互独立?

13. 已知随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}Ay\sin x, & 0\leqslant x\leqslant\dfrac{\pi}{2},0\leqslant y\leqslant x\\0, & \text{其他}\end{cases}.$$

(1) 求 A;

(2) 求边缘概率密度 $f_X(x),f_Y(y)$;

(3) X 与 Y 是否独立?

14. 设 X 和 Y 是相互独立的随机变量,X 具有概率分布律

X	1	2	3	4
p_X	0.4	0.3	0.2	0.1

Y 具有概率分布律

Y	1	2	3
p_Y	0.5	0.3	0.2

试求:(1) (X,Y) 的联合分布律;(2) $P\{X+Y\leqslant4\}$.

15. 设二维随机变量(X,Y)的联合分布律为

$X \backslash Y$	y_1	y_2	y_3
x_1	α	$\frac{1}{9}$	γ
x_2	$\frac{1}{9}$	β	$\frac{1}{3}$

若X与Y相互独立,求参数α,β,γ的值.

16. 设X和Y是相互独立的随机变量,X在区间$(0,1)$上服从均匀分布,Y服从参数为0.5的指数分布.

(1) 求(X,Y)的联合概率密度;

(2) 试求含a的二次方程$a^2+2Xa+Y=0$有实根的概率.

17. 随机变量X与Y相互独立,有相同的分布律

X	-1	0	1
p_k	0.3	0.2	0.5

试求下列函数的分布律:

(1) $Z=X+Y$;

(2) $W=\sin\dfrac{(X+Y)\pi}{2}$;

(3) $M=\max\{X,Y\}$;

(4) $N=\min\{X,Y\}$.

18. 设随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{2}(x+y)\mathrm{e}^{-(x+y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases}.$$

(1) X和Y是否相互独立?

(2) 求$Z=X+Y$的概率密度.

19. 设随机变量X在区间$(0,1)$服从均匀分布,随机变量Y具有概率密度

$$f_Y(y)=\begin{cases}3y^2, & 0<y<1\\ 0, & \text{其他}\end{cases},$$

且X,Y相互独立.求:

(1) $Z=\max\{X,Y\}$的分布函数和概率密度;

(2) $P\left\{\dfrac{1}{2}\leqslant Z<\dfrac{3}{2}\right\}$;

(3) 求$Z=X+Y$的概率密度.

20. 一工厂的工人完成某项任务所需的时间(以h计)X是一个随机变量,它具有概率密度

$$f(x)=\begin{cases}\mathrm{e}^{-(x-\theta)}, & x>\theta\\ 0, & \text{其他}\end{cases},$$

其中$\theta>0$是一常数(表示完成任务的最小时间).自这一工厂中随机地选出n个工人,分别以$X_1,X_2,\cdots,X_n$记这n个工人完成任务所需的时间,设$X_1,X_2,\cdots,X_n$相互独立.求$Z=\min\{X_1,X_2,\cdots,X_n\}$的概率密度.

21. 设 X,Y 是相互独立的随机变量,它们都服从正态分布 $N(0,\sigma^2)$. 试证随机变量 $Z=\sqrt{X^2+Y^2}$ 具有概率密度

$$f_Z(z)=\begin{cases}\dfrac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma^2}}, & z\geqslant 0\\ 0, & \text{其他}\end{cases}.$$

我们称 Z 服从参数为 $\sigma(\sigma>0)$ 的瑞利(Rayleigh)分布.

补充与提高

22. 已知 5 个元件中有两个是失效的,现进行检测,每次随机地从未检测的元件中取出一个进行检测,直到失效的都查出来为止. 记 X 表示第一个失效元件被检测出来时的检测次数,记 Y 表示其后直到第二个失效元件被检测出来时的附加次数. 求 X 和 Y 的联合分布律.

23. 设二维随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}1, & 0<x<1,0<y<2x\\ 0, & \text{其他}\end{cases}.$$

求:(1) (X,Y) 的边缘概率密度 $f_X(x),f_Y(y)$;

(2) $Z=2X-Y$ 的概率密度 $f_Z(z)$.

24. 设随机变量 X 在区间 $(0,1)$ 上服从均匀分布,在 $X=x(0<x<1)$ 的条件下,随机变量 Y 在区间 $(0,x)$ 上服从均匀分布. 求:

(1) (X,Y) 的联合概率密度;

(2) Y 的概率密度;

(3) $P\{X+Y>1\}$.

25. 设随机变量 X 与 Y 相互独立,其概率密度分别为

$$f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases} \quad \text{和} \quad f_Y(y)=\begin{cases}e^{-y} & y>0\\ 0, & y\leqslant 0\end{cases},$$

求随机变量 $Z=2X+Y$ 的概率密度.

26. 设随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}e^{-(x+y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases},$$

求 $Z=|X-Y|$ 的概率密度.

27. 设随机变量 X 和 Y 同分布,其概率密度为

$$f(x)=\begin{cases}\dfrac{3}{8}x^2, & 0<x<2\\ 0, & \text{其他}\end{cases}.$$

已知事件 $A=\{X>a\}$ 和事件 $B=\{Y>a\}$ 相互独立,且 $P(A\cup B)=\dfrac{3}{4}$,求常数 a.

28. 设 X,Y 是两个独立同分布的随机变量,分别表示两个电子元件的寿命(单位: h),其概率密度为 $f(x)=\begin{cases}\dfrac{1000}{x^2}, & x>1000\\ 0, & x\leqslant 1000\end{cases}$. 求 $Z=\dfrac{X}{Y}$ 的概率密度.

29. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y),\quad -\infty<x,y<+\infty.$$

试证：X 和 Y 皆服从 $N(0,1)$分布.

30. (二项分布的可加性)设随机变量 $X\sim B(n,p)$,$Y\sim B(m,p)$,且 X 与 Y 相互独立.试证：

$$X+Y\sim B(n+m,p).$$

$\left(\text{提示：需用到式}\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i}=\binom{n+m}{k}\right)$

31. (泊松分布的可加性)设随机变量 $X\sim P(\lambda_1)$,$Y\sim P(\lambda_2)$,且 X 与 Y 相互独立.试证：$X+Y\sim P(\lambda_1+\lambda_2)$.

第 4 章

随机变量的数字特征

从前面的讨论知道，随机变量的分布函数（分布律或概率密度）可完整地描述随机变量的概率分布情况，然而在许多实际问题中有时并不需要去全面描述随机变量，只要了解随机变量的某些特征即可. 例如，分析某班某门课程的考试成绩，主要看该班学生的平均成绩，以及每位学生的成绩与平均成绩的偏离程度. 这种“平均成绩”、“偏离程度”虽然不是对考试成绩这个随机变量的完整描述，但已反映出随机变量在某些方面的重要特征，它们在理论和实践上都具有重要意义. 我们把描述随机变量某些特征的数字称为随机变量的数字特征. 本章将介绍几个常用的数字特征：数学期望、方差、协方差、相关系数及矩.

4.1 数学期望

4.1.1 数学期望的概念

在实际工作中，人们常常使用“平均值”这个概念. 例如，为考察一批钢筋的平均抗拉强度，现从中抽取 10 根，测得抗拉强度如下：

抗拉强度 x_i	110	120	125	130	135	140
根数 n_i	1	2	3	2	1	1

则它们的平均抗拉强度为

$$\frac{\sum_{i=1}^{6} x_i n_i}{n} = \sum_{i=1}^{6} x_i \frac{n_i}{n}$$

$$= \frac{110\times1+120\times2+125\times3+130\times2+135\times1+140\times1}{10}$$

$$= 110\times\frac{1}{10}+120\times\frac{2}{10}+125\times\frac{3}{10}+130\times\frac{2}{10}+135\times\frac{1}{10}+140\times\frac{1}{10}$$

$$= 126.$$

式中 $n = \sum_{i=1}^{6} n_i$；$\frac{n_i}{n}$ 是事件“抗拉强度为 x_i”的频率，若记作 f_i，则平均抗拉强度可表示为 $\sum_{i=1}^{6} x_i f_i$.

如果另外再抽验 10 根钢筋，那么又可得到一组不同的频率，相应地，也就又可得到另一个平均抗拉强度. 每做一次这样的试验，就可得到一个平均抗拉强度. 那么这批钢筋的平均抗拉强度到底是多少呢？

由频率与概率的关系可知，在求平均值时，用概率 p_i 去代替上述和式中的频率 f_i，这样得到的平均值才是理论上的（也是真正意义上的）平均值，它不会随试验的变化而变化. 这种平均值，称为随机变量的数学期望. 其严格的数学定义如下.

定义 4.1　设离散型随机变量 X 的分布律为

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots,$$

若级数 $\sum\limits_{k=1}^{\infty}x_kp_k$ 绝对收敛，即 $\sum\limits_{k=1}^{\infty}|x_k|p_k$ 收敛，则称级数 $\sum\limits_{k=1}^{\infty}x_kp_k$ 的和为随机变量 X 的**数学期望**，记作 $E(X)$，即

$$E(X)=\sum_{k=1}^{\infty}x_kp_k. \tag{4.1}$$

若级数 $\sum\limits_{k=1}^{\infty}x_kp_k$ 不绝对收敛，则称随机变量 X 的数学期望不存在.

若 X 为连续型随机变量，其密度函数为 $f(x)$，则 X 落入 $(x_k,x_k+\mathrm{d}x)$ 内的概率可近似地表示为 $f(x_k)\mathrm{d}x$. 它与离散型随机变量的 p_k 类似，下面给出定义.

定义 4.2　设连续型随机变量 X 的密度函数为 $f(x)$，若积分 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 绝对收敛，即 $\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x$ 收敛，则称积分 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 的值为随机变量 X 的**数学期望**，记作 $E(X)$. 即

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x. \tag{4.2}$$

若积分 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 不绝对收敛，则称随机变量 X 的数学期望不存在.

数学期望简称为**期望**或**均值**，在不致混淆的情况下也可记为 EX.

随机变量 X 的数学期望 $E(X)$ 是一个常量，它是从概率的角度来计算随机变量 X 所有可能取值的平均值，具有重要的统计意义.

例 1　设甲、乙二人打靶的得分分别为随机变量 X_1,X_2，其分布律如下：

X_1	0	1	2
p_{X_1}	0.3	0.45	0.25

X_2	0	1	2
p_{X_2}	0.15	0.8	0.05

其中脱靶为 0 分，5 环以下为 1 分，6 环以上为 2 分. 问谁的水平略高一筹.

解　$E(X_1)=0\times0.3+1\times0.45+2\times0.25=0.95$，

$E(X_2)=0\times0.15+1\times0.8+2\times0.05=0.9$.

可知，甲得分的均值比乙略高，显然这只是“期望”得分.

例 2　设随机变量 X 的取值 $x_k=(-1)^k\dfrac{2^k}{k}$，$p_k=P\{X=x_k\}=\dfrac{1}{2^k}$，$k=1,2,\cdots$，讨论 X 的数学期望是否存在.

解 虽然 $\sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$ 存在，但由于 $\sum_{k=1}^{\infty} |x_k| p_k = \sum_{k=1}^{\infty} \frac{1}{k}$ 是不存在的，因此 X 的数学期望不存在.

例 3 设随机变量 X 的概率密度为 $f(x)=\begin{cases} x, & 0\leqslant x<1 \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他} \end{cases}$ 求 $E(X)$.

解
$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x \\ &= \int_{-\infty}^{0} x \times 0 \mathrm{d}x + \int_{0}^{1} x x \mathrm{d}x + \int_{1}^{2} x(2-x)\mathrm{d}x + \int_{2}^{+\infty} x \times 0 \mathrm{d}x \\ &= \frac{x^3}{3}\bigg|_0^1 + \left(x^2 - \frac{x^3}{3}\right)\bigg|_1^2 = \frac{1}{3} + \left(4-\frac{8}{3}\right) - \left(1-\frac{1}{3}\right) \\ &= 1. \end{aligned}$$

例 4 设随机变量 X 的概率密度为 $f(x)=\frac{1}{\pi(1+x^2)}, -\infty<x<+\infty$ (X 服从柯西分布)，试证 $E(X)$不存在.

解 $\int_{-\infty}^{+\infty} |x| f(x) \mathrm{d}x = \int_{-\infty}^{0} \frac{-x}{\pi(1+x^2)} \mathrm{d}x + \int_{0}^{+\infty} \frac{x}{\pi(1+x^2)} \mathrm{d}x$. 由于 $\int_{0}^{+\infty} \frac{x}{\pi(1+x^2)} \mathrm{d}x = \frac{1}{2\pi} \ln(1+x^2)\Big|_0^{+\infty} = \infty$ 不存在，从而 $\int_{-\infty}^{+\infty} |x| f(x) \mathrm{d}x$ 也不存在，所以 $E(X)$不存在.

4.1.2 随机变量函数的数学期望

在实际问题中常常需要求出随机变量函数的数学期望. 例如企业在作决策时，总是考虑平均利润，即利润的数学期望. 而利润 Y 是销售额 X 的函数 $Y=g(X)$，销售额 X 是随机变量. 这时，我们可以不必求 Y 的分布而直接由 X 的分布求 $E(Y)$. 下面的定理说明了这一点.

定理 4.1 设 Y 是随机变量 X 的函数，$Y=g(X)$，g 是连续函数.

(1) 设 X 是离散型随机变量，它的分布律为 $P\{X=x_k\}=p_k, k=1,2,\cdots$，若 $\sum_{k=1}^{\infty} g(x_k) p_k$ 绝对收敛，即 $\sum_{k=1}^{\infty} |g(x_k)| p_k$ 收敛，则

$$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k. \tag{4.3}$$

(2) 设 X 是连续型随机变量，它的概率密度为 $f(x)$，若 $\int_{-\infty}^{+\infty} g(x) f(x) \mathrm{d}x$ 绝对收敛，即 $\int_{-\infty}^{+\infty} |g(x)| f(x) \mathrm{d}x$ 收敛，则

$$E(Y) = E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) \mathrm{d}x. \tag{4.4}$$

此定理的证明超出了本书的范围，从略.

定理 4.1 还可以推广到含有两个或两个以上随机变量的函数的情况，这里我们给出二维随机变量的函数 $Z=g(X,Y)$的情况.

定理 4.2 设 Z 是随机变量 X,Y 的函数,$Z=g(X,Y)$,g 为连续函数,则 Z 是一维随机变量.

(1) 若(X,Y)是离散型随机变量,它的联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

则

$$E(Z)=E[g(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}g(x_i,y_j)p_{ij}. \tag{4.5}$$

这里,设上式右端的级数绝对收敛.

(2) 若(X,Y)是连续型随机变量,它的概率密度为 $f(x,y)$,则

$$E(Z)=E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y. \tag{4.6}$$

这里,设上式右端的积分绝对收敛.

例 5 设随机变量 X 的分布律如下:

X	-2	0	2
p_k	0.4	0.3	0.3

求:$E(X),E(X^2),E(3X^2+5)$.

解 $E(X)=-2\times0.4+0\times0.3+2\times0.3=-0.2$,

$E(X^2)=(-2)^2\times0.4+0^2\times0.3+2^2\times0.3=2.8$,

$E(3X^2+5)=[3\times(-2)^2+5]\times0.4+(3\times0^2+5)\times0.3+$
$(3\times2^2+5)\times0.3=1.34.$

例 6 某商店一天的营业额 X(以万元计)是一个随机变量.已知 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{6}{181}(11x-x^2), & 5<x<6\\ 0, & \text{其他}\end{cases},$$

一天的盈利 $Y=g(X)=\dfrac{1}{2}X-1$.求 $E(Y)$.

解 $$E(Y)=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x=\int_5^6\left(\frac{1}{2}x-1\right)\frac{6}{181}(11x-x^2)\mathrm{d}x$$
$$=1.75(\text{万元}).$$

例 7 某公司经销某种原料,根据历史资料可知这种原料的市场需求量 X(单位:t)服从(300,500)上的均匀分布.每售出 1 t 该原料,公司可获利 1.5 千元;若积压 1 t,则公司损失 0.5 千元.问公司应该组织多少货源,可使平均收益最大.

解 设公司组织该货源 at(a 待定),则应该有 $300\leqslant a\leqslant500$,又设 Y 为收益,则 Y 是 X 的函数,具体为

$$Y=g(X)=\begin{cases}1.5a, & X\geqslant a\\ 1.5X-0.5(a-X)=2X-0.5a, & X<a\end{cases},$$

$$X\text{ 的概率密度 } f_X(x)=\begin{cases}\dfrac{1}{200}, & 300<x<500\\ 0, & \text{其他}\end{cases}.$$

于是

$$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty}g(x)f_X(x)\mathrm{d}x=\int_{300}^{500}g(x)\times\frac{1}{200}\mathrm{d}x$$
$$=\frac{1}{200}\left[\int_{300}^{a}(2x-0.5a)\mathrm{d}x+\int_{a}^{500}1.5a\mathrm{d}x\right]$$
$$=\frac{1}{200}(-a^2+900a-300^2).$$

令$\frac{\mathrm{d}E(Y)}{\mathrm{d}a}=\frac{1}{200}(-2a+900)=0$，得 $a=450$，而

$$\frac{\mathrm{d}^2E(Y)}{\mathrm{d}a^2}=-\frac{1}{100}<0,$$

所以组织货源 450 t 时，公司平均收益最大.

例 8 设二维随机变量(X,Y)的分布律为

X \ Y	0	1	2	3
1	0	$\frac{3}{8}$	$\frac{3}{8}$	0
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$

求：(1)$E(X)$；(2)$E(Y)$；(3)$E(XY)$.

解 (1) X 的边缘分布律为

X	1	3
p_X	$\frac{3}{4}$	$\frac{1}{4}$

$$E(X)=1\times\frac{3}{4}+3\times\frac{1}{4}=\frac{3}{2};$$

(2) Y 的边缘分布律为

Y	0	1	2	3
p_Y	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$E(Y)=0\times\frac{1}{8}+1\times\frac{3}{8}+2\times\frac{3}{8}+3\times\frac{1}{8}=\frac{3}{2};$$

(3) $$E(XY)=(1\times0)\times0+(1\times1)\times\frac{3}{8}+(1\times2)\times\frac{3}{8}+(1\times3)\times0+$$
$$(3\times0)\times\frac{1}{8}+(3\times1)\times0+(3\times2)\times0+(3\times3)\times\frac{1}{8}$$
$$=\frac{9}{4}.$$

例 9 设二维随机变量(X,Y)的概率密度为 $f(x,y)=\begin{cases}12y^2, & 0\leqslant y\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}$. 求：(1)$E(X)$；(2)$E(Y)$；(3)$E(XY)$.

解 (1) $E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^x x\cdot 12y^2\mathrm{d}y=\int_0^1 x\cdot 4y^3\Big|_0^x\mathrm{d}x=\frac{4}{5}$;

(2) $E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^x y\cdot 12y^2\mathrm{d}y=\int_0^1 3y^4\Big|_0^x\mathrm{d}x=\frac{3}{5}$;

(3) $E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^x xy12y^2\mathrm{d}y=\int_0^1 x\cdot 3y^4\Big|_0^x\mathrm{d}x=\frac{1}{2}$.

注：另可先求出 X,Y 的边缘概率密度，再求 $E(X),E(Y)$.

4.1.3 数学期望的性质

数学期望具有如下性质(设以下所遇到的随机变量的数学期望都存在)：

(1) 设 c 是常数，则 $E(c)=c$.

(2) 设 X 是一个随机变量，c 是常数，则 $E(cX)=cE(X)$.

(3) 设 X,Y 是任意两个随机变量，则 $E(X+Y)=E(X)+E(Y)$.

性质 3 可推广到任意有限个随机变量之和的情况.

综合性质 1,2,3 有结论 $E(aX+b)=aE(X)+b$.

(4) 设 X,Y 是两个相互独立的随机变量，则 $E(XY)=E(X)E(Y)$.

性质 4 可推广到任意有限个相互独立的随机变量之积的情况.

性质 1、性质 2 请读者自己完成证明. 下面以连续型随机变量为例，给出性质 3、性质 4 的证明，离散型的情况类似可证.

证明性质 3：只就 X 为连续型随机变量的情况来证明，对于离散型的情况，其证明类似.

设二维随机变量 (X,Y) 的概率密度为 $f(x,y)$，则有

$$\begin{aligned}E(X+Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y\\&=E(X)+E(Y).\end{aligned}$$

证明性质 4：只就 X 为连续型随机变量的情况来证明，对于离散型的情况，其证明类似.

设二维随机变量 (X,Y) 的概率密度为 $f(x,y)$，由于 X 与 Y 相互独立，$f(x,y)=f_X(x)\cdot f_Y(y)$，其中 $f_X(x)$ 与 $f_Y(y)$ 分别是 (X,Y) 关于 X 和关于 Y 的边缘概率密度，则有

$$\begin{aligned}E(XY)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(xy)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x\int_{-\infty}^{+\infty}yf_Y(y)\mathrm{d}y=E(X)E(Y).\end{aligned}$$

数学期望的良好性质有助于简化计算.

例 10 (4.1 节例 5 续)设随机变量 X 的分布律如下：

X	-2	0	2
p_k	0.4	0.3	0.3

求 $E(X),E(2X),E(3X-7)$.

解 由期望的性质有

$$E(X)=-2\times0.4+0\times0.3+2\times0.3=-0.2,$$
$$E(2X)=2E(X)=2\times(-0.2)=-0.4,$$
$$E(3X-7)=3E(X)-7=-7.6.$$

例 11 一民航的送客车载有 20 位旅客,自机场开出,沿途旅客可由 10 个车站下车.如到达一个车站没有旅客下车就不停车,设每位旅客在各个车站下车是等可能的,且各旅客是否下车相互独立,以 X 表示停车的次数.求 $E(X)$.

解 设随机变量 $X_i=\begin{cases}1, & 第\ i\ 个车站有人下车\\ 0, & 第\ i\ 个车站无人下车\end{cases}$,$i=1,2,\cdots,10$,则

$$X=X_1+X_2+\cdots+X_{10}.$$

由题意:任一旅客在第 i 个车站不下车的概率为 $\frac{9}{10}$,20 位旅客都不在第 i 站下车的概率为 $\left(\frac{9}{10}\right)^{20}$,故 $P\{X_i=0\}=\left(\frac{9}{10}\right)^{20}$,$P\{X_i=1\}=1-\left(\frac{9}{10}\right)^{20}$.于是,$X_i$ 的分布律为

X_i	0	1
p_i	$\left(\frac{9}{10}\right)^{20}$	$1-\left(\frac{9}{10}\right)^{20}$

因此

$$E(X_i)=0\times\left(\frac{9}{10}\right)^{20}+1\times\left[1-\left(\frac{9}{10}\right)^{20}\right]=1-\left(\frac{9}{10}\right)^{20},\quad i=1,2,\cdots,10.$$

从而得

$$E(X)=E(X_1+X_2+\cdots+X_{10})=E(X_1)+E(X_2)+\cdots+E(X_{10})$$
$$=10\times\left[1-\left(\frac{9}{10}\right)^{20}\right]\approx 8.8.$$

这表明班车平均停车约 9 次.

类似本例将 X 分解为若干随机变量之和,然后用数学期望的性质再求 $E(X)$的方法,具有一定的普遍意义,若使用得当,可使复杂的问题简单化.

例 12 设一电路中电流 I(单位:A)与电阻 R(单位:Ω)是两个相互独立的随机变量,其概率密度分别为

$$f_I(i)=\begin{cases}2i, & 0\leqslant i\leqslant 1\\ 0, & 其他\end{cases},\quad f_R(r)=\begin{cases}\frac{r^2}{9}, & 0\leqslant r\leqslant 3\\ 0, & 其他\end{cases}.$$

试求电压(单位:V)$V=IR$ 的均值.

解 $$E(V)=E(IR)=E(I)E(R)=\left(\int_{-\infty}^{+\infty}if_I(i)\mathrm{d}i\right)\left(\int_{-\infty}^{+\infty}rf_R(r)\mathrm{d}r\right)$$
$$=\left(\int_0^1 2i^2\mathrm{d}i\right)\left(\int_0^3\frac{r^3}{9}\mathrm{d}r\right)=\frac{3}{2}\mathrm{V}.$$

4.2　方　　差

数学期望反映了随机变量取值的平均程度，是一个很重要的数字特征. 但是在某些场合只知道平均程度是不够的. 例如包装机包装食盐，额定重量每袋为 500 g，实际上不可能做到每袋重量正好是额定重量，每袋盐的重量是随机重量. 现有两台机器，包装出的食盐重量都保证了期望值为 500 g，但是一台机器包装的重量或者重了约 100 g，或者轻了约 100 g；另一台机器包装的重量或者重了约 2 g，或者轻了约 2 g. 比较两台机器，显然后一台机器的性能较好，因为后者与期望值的偏离程度要小. 为了刻划随机变量取值与其数学期望值的偏离程度，容易想到用 $E[X-E(X)]$来表示. 但由于 X 是一个随机变量，随机变量 X 与 $E(X)$的偏离量 $X-E(X)$也是一个随机变量，平均来讲，正、负偏离量相互抵消，因而 $E[X-E(X)]=0$，故不能用 $E[X-E(X)]$来考虑平均偏离量. 为消除符号的影响，最简单的是考虑偏离量的绝对值的平均，即 $E[|X-E(X)|]$；又由于绝对值运算在数学上处理很不方便，所以我们采用 $E\{[X-E(X)]^2\}$作为描述随机变量与其均值的偏离程度的数字特征，并且我们将这个数字特征称为随机变量 X 的方差.

4.2.1　方差的定义

定义 4.3　设 X 为随机变量，如果 $E\{[X-E(X)]^2\}$存在，则称其为 X 的**方差**，记作 $D(X)$，即$D(X)=E\{[X-E(X)]^2\}$. 并称 $\sqrt{D(X)}$为 X 的**标准差**或**均方差**，记作 $\sigma(X)$.

由定义可知，方差实际上就是随机变量 X 的函数$[X-E(X)]^2$ 的数学期望，于是可得以下结论.

(1) 若 X 是离散型随机变量，分布律为 $P\{X=x_k\}=p_k$，$k=1,2,\cdots$，则

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2 p_k. \tag{4.7}$$

(2) 若 X 是连续型随机变量，概率密度为 $f(x)$，则

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)\mathrm{d}x. \tag{4.8}$$

关于方差的计算，在实际中常用下面的简化公式：

$$D(X)=E(X^2)-[E(X)]^2. \tag{4.9}$$

证明

$$\begin{aligned}D(X)&=E\{[X-E(X)]^2\}=E\{X^2-2XE(X)+[E(X)]^2\}\\&=E(X^2)-2E(X)E(X)+[E(X)]^2\\&=E(X^2)-[E(X)]^2.\end{aligned}$$

例 1(4.1 节例 5 续)　设随机变量 X 的分布律为

X	-2	0	2
p_k	0.4	0.3	0.3

求 $D(X)$.

解 已算得 $E(X)=-0.2, E(X^2)=2.8$,于是,$D(X)=E(X^2)-[E(X)]^2=2.8-(-0.2)^2=2.76$.

例 2(4.1 节例 3 续) 设随机变量 X 的概率密度为 $f(x)=\begin{cases}x, & 0\leqslant x<1\\ 2-x, & 1\leqslant x<2\\ 0, & \text{其他}\end{cases}$. 求 $D(X)$.

解 已算得

$$E(X)=1,$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2 f(x)\mathrm{d}x=\int_0^1 x^2\cdot x\mathrm{d}x+\int_1^2 x^2(2-x)\mathrm{d}x=\frac{7}{6}.$$

于是

$$D(X)=E(X^2)-[E(X)]^2=\frac{7}{6}-1=\frac{1}{6}.$$

4.2.2 方差的性质

方差具有如下性质(设以下所遇到的随机变量的方差都存在):

(1) 设 c 是常数,则 $D(c)=0$.

(2) 设 c 是常数,则 $D(cX)=c^2D(X)$.

性质 1,2 由读者自己完成证明.

(3) 若随机变量 X,Y 相互独立,则 $D(X\pm Y)=D(X)+D(Y)$.

证明

$$\begin{aligned}D(X\pm Y)&=E\{[(X\pm Y)-E(X\pm Y)]^2\}\\&=E\{[(X-E(X))\pm(Y-E(Y))]^2\}\\&=E\{[(X-E(X))]^2\}+E\{[(Y-E(Y))]^2\}\pm\\&\quad 2E\{[X-E(X)][Y-E(Y)]\}\\&=D(X)+D(Y)\pm 2E[XY-XE(Y)-YE(X)+E(X)E(Y)]\\&=D(X)+D(Y)\pm 2[E(XY)-E(X)E(Y)]\\&=D(X)+D(Y).\end{aligned}$$

性质 3 可推广到 n 个随机变量的场合.若随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,则

$$D\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n D(X_i).$$

综合性质 1,2,3 有结论

$$D(aX+b)=a^2D(X).$$

(4) $D(X)=0$ 的充要条件是 X 以概率 1 取常数 c,即 $P\{X=c\}=1$,其中 $c=E(X)$.(证明略)

例 3 设随机变量 X,Y 相互独立,且 $D(X)=\sigma_1^2, D(Y)=\sigma_2^2$,试求 $D(2X-5Y+4)$.

解 $D(2X-5Y+4)=D(2X-5Y)=2^2D(X)+5^2D(Y)=4\sigma_1^2+25\sigma_2^2$.

例 4 设随机变量 X 具有期望 $E(X)=\mu$,方差 $D(X)=\sigma^2\neq 0$. 设 $X^*=\dfrac{X-\mu}{\sigma}$,求

$E(X^*)$，$D(X^*)$.

解 $$E(X^*) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu) = \frac{1}{\sigma}(\mu-\mu) = 0,$$

$$D(X^*) = D\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}D(X) = \frac{1}{\sigma^2}D(X) = \frac{1}{\sigma^2}\sigma^2 = 1.$$

称 $X^* = \frac{X-\mu}{\sigma}$ 为随机变量 X 的**标准化**. 该题说明对任意随机变量 X，若期望、方差存在且方差不为0，标准化后得到的随机变量 X^* 都有 $E(X^*)=0$，$D(X^*)=1$.

4.3 几种常用分布的期望、方差

本节主要介绍几种常用分布的期望和方差(见表4-1).

表 4-1

分 布	参 数	分布律或概率密度	期 望	方 差
(0-1)分布	$0<p<1$	$P\{X=k\}=p^k(1-p)^{1-k}, k=0,1$	p	pq
二项分布	$n\geqslant 1$ $0<p<1$	$P\{X=k\}=C_n^k p^k(1-p)^{n-k}$, $k=0,1,\cdots,n$	np	npq
泊松分布	$\lambda>0$	$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}$, $k=0,1,2,\cdots$	λ	λ
均匀分布	$a<b$	$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$\theta>0$	$f(x)=\begin{cases}\theta e^{-\theta x} & x>0\\ 0, & x\leqslant 0\end{cases}$	$\frac{1}{\theta}$	$\frac{1}{\theta^2}$
正态分布	μ $\sigma>0$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

1. (0-1) 分布

设随机变量 X 的分布律为

X	1	0
p_k	p	q

其中 $0<p<1$，$q=1-p$，则 $E(X)=p$，$D(X)=pq$.

证明 显然

$$E(X) = p, E(X^2) = 1^2\times p + 0^2\times q = p,$$

$$D(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p) = pq.$$

2. 二项分布

设随机变量 $X\sim B(n,p)$，其分布律为 $P\{X=k\}=C_n^k p^k q^{n-k}$($k=0,1,2,\cdots,n$；$0<p<1$，

$q=1-p$),则 $E(X)=np, D(X)=npq$.

证明 在 n 重伯努利试验中,每次试验事件 A 发生的概率为 p,不发生的概率为 $q=1-p$. 若引入随机变量

$$X_i=\begin{cases}1, & \text{第 } i \text{ 次试验 } A \text{ 发生} \\ 0, & \text{第 } i \text{ 次试验 } A \text{ 不发生}\end{cases}, \quad i=1,2,\cdots,n,$$

则 A 发生的次数 $X=X_1+X_2+\cdots+X_n$, $X\sim B(n,p)$, $X_i\sim(0\text{-}1)$分布,且 $X_1,X_2,\cdots,X_n$ 是相互独立的. 于是

$$\begin{aligned}E(X)&=E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)\\&=nE(X_i)=np,\\D(X)&=D(X_1+X_2+\cdots+X_n)=D(X_1)+D(X_2)+\cdots+D(X_n)\\&=nD(X_i)=npq.\end{aligned}$$

3. 泊松分布

设随机变量 $X\sim P(\lambda)$,分布律为 $P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}$ ($k=0,1,2,\cdots;\lambda>0$ 是常数),则 $E(X)=\lambda, D(X)=\lambda$.

证明

$$\begin{aligned}E(X)&=\sum_{k=0}^{\infty}k\frac{\lambda^k e^{-\lambda}}{k!}=\sum_{k=1}^{\infty}\frac{\lambda^k e^{-\lambda}}{(k-1)!}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\\&=\lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda,\\D(X)&=E(X^2)-[E(X)]^2=E[X(X-1)+X]-\lambda^2\\&=E[X(X-1)]+\lambda-\lambda^2=\sum_{k=1}^{\infty}k(k-1)\frac{\lambda^k}{k!}e^{-\lambda}+\lambda-\lambda^2\\&=\lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda-\lambda^2\\&=\lambda^2 e^{-\lambda}e^{\lambda}+\lambda-\lambda^2=\lambda.\end{aligned}$$

4. 均匀分布

设随机变量 $X\sim U(a,b)$,概率密度为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases},$$

则

$$E(X)=\frac{a+b}{2},\quad D(X)=\frac{(b-a)^2}{12}.$$

证明

$$E(X)=\int_a^b x\cdot\frac{1}{b-a}dx=\frac{1}{b-a}\cdot\frac{x^2}{2}\Big|_a^b=\frac{a+b}{2},$$

$$D(X)=E(X^2)-[E(X)]^2=\int_a^b x^2\cdot\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2$$

$$= \frac{1}{b-a} \cdot \frac{x^3}{3} \Big|_a^b - \left(\frac{a+b}{2}\right)^2$$
$$= \frac{(b-a)^2}{12}.$$

5. 指数分布

设随机变量 $X \sim \mathrm{Exp}(\theta)$，概率密度为

$$f(x) = \begin{cases} \theta e^{-\theta x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}, \quad \theta > 0 \text{ 为常数},$$

则

$$E(X) = \frac{1}{\theta}, \quad D(X) = \frac{1}{\theta^2}.$$

证明

$$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x = \int_0^{+\infty} x\theta e^{-\theta x} \mathrm{d}x = -\int_0^{+\infty} x \mathrm{d}e^{-\theta x}$$
$$= -x e^{-\theta x} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\theta x} \mathrm{d}x = -\frac{1}{\theta} e^{-\theta x} \Big|_0^{+\infty} = \frac{1}{\theta},$$
$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) \mathrm{d}x = \int_0^{+\infty} x^2 \theta e^{-\theta x} \mathrm{d}x = -\int_0^{+\infty} x^2 \mathrm{d}e^{-\theta x}$$
$$= -x^2 e^{-\theta x} \Big|_0^{+\infty} + \int_0^{+\infty} 2x e^{-\theta x} \mathrm{d}x = -\frac{2}{\theta} \int_0^{+\infty} x \mathrm{d}e^{-\theta x}$$
$$= -\frac{2}{\theta} x e^{-\theta x} \Big|_0^{+\infty} + \frac{2}{\theta} \int_0^{+\infty} e^{-\theta x} \mathrm{d}x = -\frac{2}{\theta^2} e^{-\theta x} \Big|_0^{+\infty} = \frac{2}{\theta^2},$$
$$D(X) = E(X^2) - [E(X)]^2 = \frac{2}{\theta^2} - \frac{1}{\theta^2} = \frac{1}{\theta^2}.$$

6. 正态分布

设随机变量 $X \sim N(\mu, \sigma^2)$，概率密度为 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$-\infty < x < +\infty$；$\mu, \sigma(\sigma > 0)$ 为常数，则 $E(X) = \mu$，$D(X) = \sigma^2$.

证明

$$E(X) = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x + \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x$$
$$\xlongequal{\text{令 } t = x-\mu} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2\sigma^2}} \mathrm{d}t + \mu \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \mathrm{d}t = \mu,$$
$$D(X) = E[X - E(X)]^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x$$
$$\xlongequal{\text{令 } t = \frac{x-\mu}{\sigma}} \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} \mathrm{d}t = \frac{\sigma^2}{\sqrt{2\pi}} \left[t(-e^{-\frac{t^2}{2}}) \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} \mathrm{d}t \right]$$
$$= \frac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = \sigma^2.$$

4.4 协方差与相关系数 矩

对于二维随机变量(X,Y)，我们除了讨论X与Y的数学期望和方差以外，还需要讨论描述X与Y之间相互关系的数字特征：协方差与相关系数.

4.4.1 协方差

由方差性质3的证明我们知道，当X与Y相互独立时，$E\{[X-E(X)][Y-E(Y)]\}=0$，这意味着若$E\{[X-E(X)][Y-E(Y)]\}\neq 0$，则$X$与$Y$肯定不是相互独立的. 这说明$E\{[X-E(X)][Y-E(Y)]\}$的数值在一定程度上反映了$X$与$Y$之间的相互关系. 因而我们引入以下定义.

定义 4.4 若$E\{[X-E(X)][Y-E(Y)]\}$存在，则称之为随机变量X与Y的**协方差**，记为$\mathrm{Cov}(X,Y)$，即

$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}. \tag{4.10}$$

由方差性质3的证明可知，对于任意两个随机变量X与Y，有

$$D(X\pm Y)=D(X)+D(Y)\pm 2\mathrm{Cov}(X,Y). \tag{4.11}$$

计算协方差的公式为

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y). \tag{4.12}$$

协方差具有下列性质：

(1) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$；

(2) $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$；

(3) $\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y)$.

以上性质可由定义4.4直接推出，请读者自己完成证明.

4.4.2 相关系数

定义 4.5 设(X,Y)为二维随机变量，若$D(X)>0,D(Y)>0$，则称

$$\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为随机变量X与Y的**相关系数**，记作ρ_{XY}，即

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}. \tag{4.13}$$

相关系数ρ_{XY}与协方差$\mathrm{Cov}(X,Y)$在数值上只相差一个倍数，但相关系数是无量纲的绝对量，不受所用度量单位的影响，这样能更好地反映X与Y之间的关系. 实际上，相关系数ρ_{XY}是标准化了的随机变量X^*与Y^*的协方差. 即$\rho_{XY}=\mathrm{Cov}(X^*,Y^*)\left(其中 X^*=\dfrac{X-E(X)}{\sqrt{D(X)}},\right.$ $\left.E(X^*)=0,D(X^*)=1\right)$.

例 1 设二维随机变量(X,Y)的概率密度为$f(x,y)=\begin{cases}2, & x>0,y>0,x+y<1,\\ 0, & 其他,\end{cases}$求

$\mathrm{Cov}(X,Y),\rho_{XY},D(X+Y)$.

解 $E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1 x\mathrm{d}x\int_0^{1-x}\mathrm{d}y=2\int_0^1 x(1-x)\mathrm{d}x=\frac{1}{3}$;

$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1\mathrm{d}x\int_0^{1-x}y\mathrm{d}y=\int_0^1(1-x)^2\mathrm{d}x=\frac{1}{3}$;

$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1 x\mathrm{d}x\int_0^{1-x}y\mathrm{d}y=\int_0^1 x(1-x)^2\mathrm{d}x=\frac{1}{12}$;

$E(X^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^2f(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1 x^2\mathrm{d}x\int_0^{1-x}\mathrm{d}y=2\int_0^1 x^2(1-x)\mathrm{d}x=\frac{1}{6}$;

$E(Y^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y^2f(x,y)\mathrm{d}x\mathrm{d}y=2\int_0^1\mathrm{d}x\int_0^{1-x}y^2\mathrm{d}y=\frac{2}{3}\int_0^1(1-x)^3\mathrm{d}x=\frac{1}{6}$;

$D(X)=E(X^2)-[E(X)]^2=\frac{1}{6}-\frac{1}{9}=\frac{1}{18}$;

$D(Y)=E(Y^2)-[E(Y)]^2=\frac{1}{6}-\frac{1}{9}=\frac{1}{18}$.

于是

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{12}-\frac{1}{3}\times\frac{1}{3}=-\frac{1}{36},$$

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=-\frac{\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}}=-\frac{1}{2},$$

$$D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X,Y)=\frac{1}{18}+\frac{1}{18}+2\times\left(-\frac{1}{36}\right)=\frac{1}{18}.$$

下面我们来讨论相关系数的性质，进而说明相关系数反映了随机变量间的一种相互关系的本质.

相关系数具有以下性质：

(1) $|\rho_{XY}|\leqslant 1$.

证明 由

$$\begin{aligned}D(X^*\pm Y^*)&=D(X^*)+D(Y^*)\pm 2\mathrm{Cov}(X^*,Y^*)\\&=1+1\pm 2\rho_{XY}=2(1\pm\rho_{XY})\geqslant 0,\end{aligned}$$

得

$$1\pm\rho_{XY}\geqslant 0.$$

即

$$|\rho_{XY}|\leqslant 1.$$

(2) $|\rho_{XY}|=1$ 的充要条件是 X 和 Y 以概率 1 存在线性关系，即 $P\{Y=aX+b\}=1(a\neq 0,b$ 是常数).

证明 ① 充分性

设 $Y=aX+b$，则

$$E(Y)=aE(X)+b,D(Y)=a^2D(X),$$
$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$

$$= E\{[X-E(X)][aX+b-aE(X)-b]\}$$
$$= aE[X-E(X)]^2 = aD(X),$$
$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{aD(X)}{\sqrt{D(X)}\sqrt{a^2D(X)}} = \frac{a}{|a|}.$$

当 $a>0$ 时，$\rho_{XY}=1$；当 $a<0$ 时，$\rho_{XY}=-1$.

② 必要性

设 $\rho_{XY}=1$.

由性质 1 的证明知

$$D(X^*-Y^*)=0.$$

由方差的性质 4 知 $P\{X^*-Y^*=C\}=1$，且 $C=E(X^*-Y^*)=0$，即

$$P\{X^*=Y^*\}=1.$$

由此可得 $P\{Y=aX+b\}=1$，其中 $a=\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}$，$b=E(Y)-\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}E(X)$. 此时称 X 与 Y **正相关**.

设 $\rho_{XY}=-1$，同理可证 $P(Y=aX+b)=1$，其中 $a=-\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}$，$b=E(Y)+\frac{\sqrt{D(Y)}}{\sqrt{D(X)}}\cdot E(X)$. 此时称 X 与 Y **负相关**.

由此可知，随机变量的相关系数实质上只表示随机变量之间的线性相关性. 当 $|\rho_{XY}|$ 变小时，X 和 Y 的线性相关程度就变弱，如果 $\rho_{XY}=0$，X 和 Y 之间就不存在线性关系.

定义 4.6 若随机变量 X 与 Y 的相关系数为零，则称 X 与 Y **不相关**.

为进一步研究 X 与 Y 的不相关性，我们引入下面两个定理.

定理 4.3 若 X 与 Y 相互独立，则 X 与 Y 不相关.

证明 仅对连续型随机变量情形进行证明，离散型随机变量情形请读者自己完成证明. 因为 X 与 Y 独立，所以有

$$f(x,y)=f_X(x)f_Y(y),$$

则

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]f(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\int_{-\infty}^{+\infty}[x-E(X)]f_X(x)\mathrm{d}x\int_{-\infty}^{+\infty}[y-E(Y)]f_Y(y)\mathrm{d}y$$
$$=E\{[X-E(X)][Y-E(Y)]\}=0,$$

得 X 与 Y 不相关.

定理 4.4 对随机变量 X 与 Y，下面的事实是等价的：

(1) $\mathrm{Cov}(X,Y)=0$；

(2) X 与 Y 不相关；

(3) $E(XY)=E(X)E(Y)$；

(4) $D(X+Y)=D(X)+D(Y)$.

请读者自己完成证明.

例 2 已知离散型随机变量 X 的分布律为

X	-1	0	1
p_k	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

令 $Y=X^2$,则 X 与 Y 是否不相关? X 与 Y 是否独立?

解 $E(X)=-1\times\frac{1}{3}+0\times\frac{1}{3}+1\times\frac{1}{3}=0,$

$$E(Y)=E(X^2)=(-1)^2\times\frac{1}{3}+0^2\times\frac{1}{3}+1^2\times\frac{1}{3}=\frac{2}{3},$$

$$E(XY)=E(X^3)=(-1)^3\times\frac{1}{3}+0^3\times\frac{1}{3}+1^3\times\frac{1}{3}=0,$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0.$$

说明 X 与 Y 不相关.但 X 与 Y 有函数关系 $Y=X^2$,说明 X 与 Y 不独立.

由以上讨论知道,“X 和 Y 不相关”与“X 和 Y 相互独立”是两个不同的概念. “不相关”只说明 X 和 Y 不存在线性关系,而“相互独立”说明 X 和 Y 之间既不存在线性关系,也不存在其他关系. 所以“相互独立”必导致“不相关”;反之,“不相关”可能是“相互独立”的,也可能是“不相互独立”的.

例 3 设随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. 求 ρ_{XY},并说明对二维正态随机变量来说,X 与 Y 相互独立$\Leftrightarrow$$X$ 与 Y 不相关.

解 在 3.2 节例 5 中已经知道 (X,Y) 的边缘概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},\quad -\infty<x<+\infty;$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}},\quad -\infty<y<+\infty.$$

故知 $E(X)=\mu_1,D(X)=\sigma_1^2,E(Y)=\mu_2,D(Y)=\sigma_2^2$,而

$$\begin{aligned}\mathrm{Cov}(X,Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(x-\mu_2)f(x,y)\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\cdot\\&\quad\exp^{-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right\}}\mathrm{d}x\mathrm{d}y.\end{aligned}$$

令

$$t=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right),\quad u=\frac{x-\mu_1}{\sigma_1},$$

则有

$$\begin{aligned}\mathrm{Cov}(X,Y)&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\sigma_1\sigma_2\sqrt{1-\rho^2}tu+\rho\sigma_1\sigma_2u^2)e^{-\frac{u^2}{2}-\frac{t^2}{2}}\mathrm{d}t\mathrm{d}u\\&=\frac{\rho\sigma_1\sigma_2}{2\pi}\left(\int_{-\infty}^{+\infty}u^2e^{-\frac{u^2}{2}}\mathrm{d}u\right)\cdot\left(\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\mathrm{d}t\right)+\\&\quad\frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\left(\int_{-\infty}^{+\infty}ue^{-\frac{u^2}{2}}\mathrm{d}u\right)\cdot\left(\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\mathrm{d}t\right)\end{aligned}$$

$$=\frac{\rho\sigma_1\sigma_2}{2\pi}\cdot\sqrt{2\pi}\cdot\sqrt{2\pi}=\rho\sigma_1\sigma_2,$$

于是

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sigma_1\sigma_2}=\frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2}=\rho.$$

这就是说，二维正态随机变量(X,Y)的概率密度中的参数ρ就是X与Y的相关系数，因而二维正态随机变量的分布完全可由X,Y各自的数学期望、方差以及它们的相关系数所确定.

在3.4节例4中，我们证明了X与Y相互独立的充要条件是$\rho=0$，现在知道$\rho_{XY}=\rho$. 故对于二维正态随机变量(X,Y)来说，X与Y相互独立与X,Y不相关是等价的.

4.4.3　矩与协方差矩阵

数学期望、方差和协方差都是随机变量最常用的数字特征，把它们推广到更一般的情况就有如下矩的定义. 应该指出的是，矩是统计学中最基本也是最重要的概念之一.

定义 4.7　设X与Y是随机变量，则有以下定义：

(1) 若$E(X^k)$ $(k=1,2,\cdots)$存在，称之为X的**k阶原点矩**，简称**k阶矩.**

(2) 若$E\{[X-E(X)]^k\}$ $(k=2,3,\cdots)$存在，称之为X的**k阶中心矩.**

(3) 若$E(X^kY^l)$ $(k,l=1,2,\cdots)$存在，称之为X和Y的**$k+l$阶混合矩.**

(4) 若$E\{[X-E(X)]^k[Y-E(Y)]^l\}(k,l=1,2,\cdots)$存在，称之为$X$和$Y$的**$k+l$阶混合中心矩.**

显然X的数学期望$E(X)$是X的一阶原点矩，方差$D(X)$是X的二阶中心矩，协方差$\mathrm{Cov}(X,Y)$是X和Y的二阶混合中心矩.

定义 4.8　设n维随机变量$(X_1,X_2,\cdots,X_n)$的二阶混合中心矩

$$C_{ij}=\mathrm{Cov}(X_i,X_j)=E\{[X_i-E(X_i)][X_j-E(X_j)]\}$$

都存在，则矩阵

$$\boldsymbol{C}=\begin{bmatrix}C_{11}&C_{12}&\cdots&C_{1n}\\C_{21}&C_{22}&\cdots&C_{2n}\\\vdots&\vdots&&\vdots\\C_{n1}&C_{n2}&\cdots&C_{nn}\end{bmatrix}$$

称为$(X_1,X_2,\cdots,X_n)$的**协方差矩阵.**

由于$C_{ii}=D(X_i),C_{ij}=\mathrm{Cov}(X_i,X_j)=\mathrm{Cov}(X_j,X_i)=C_{ji},i\neq j,i,j=1,2,\cdots,n$，所以协方差矩阵是一个对称矩阵.

小　结

随机变量的数字特征是概率分布的某种表征，是描述随机变量特征的有效工具. 特别是重要的几个数字特征：数学期望、方差、相关系数等都有明确的概率意义，同时又具有良好的性质. 因此数字特征的概念在概率论与数理统计中具有很重要的地位.

求随机变量的数字特征，实际上多可归结为求随机变量及其函数的数学期望问题. 而

某些随机变量的数学期望如用定义求会比较麻烦，甚至难以计算，如果利用好数学期望的性质，求解可能变得更加容易.

基本要求：

(1) 理解随机变量的数学期望与方差的概念，熟练掌握数学期望与方差的性质，从而运用它们来进行相关的运算；掌握根据随机变量的概率分布求其函数的数学期望.

(2) 熟记常用分布((0-1)分布、二项分布、泊松分布、均匀分布、指数分布和正态分布)的数学期望与方差.

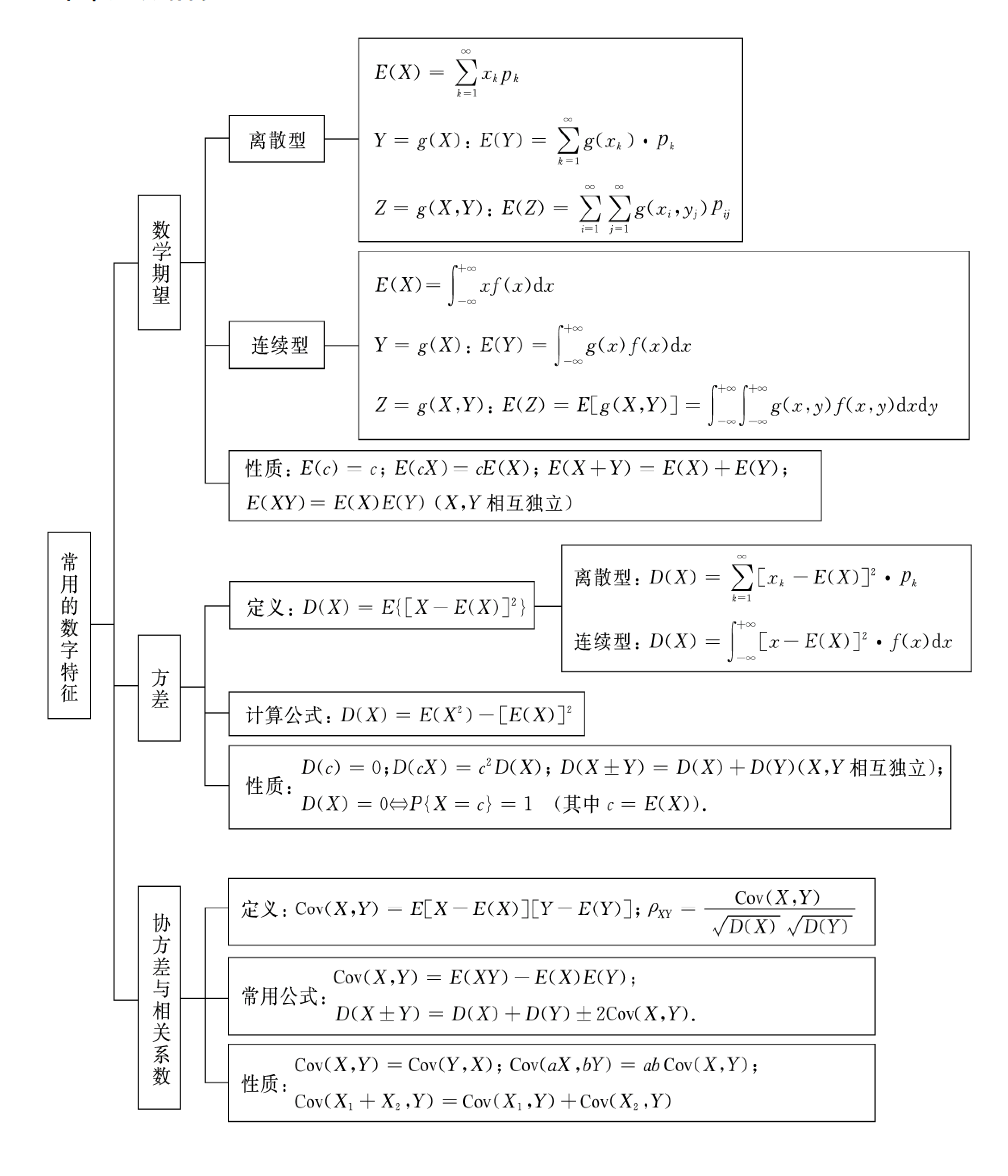
(3) 了解协方差与相关系数的概念与性质，并会运用它们计算具体分布的协方差与相关系数.

(4) 了解矩与协方差矩阵的概念.

重点：数学期望与方差的概念、性质以及相关计算.

难点：相互独立和不相关之间的区别与联系.

本章知识结构：



习　题　4

1. (1) 设 X,Y 为两个相互独立的随机变量,且 $E(X)=E(Y)=1,D(X)=D(Y)=1$,则 $E(X+Y)=$________,$D(X-Y)=$________.

(2) 对随机变量 X,Y 有 $D(X)=4,D(Y)=9,\rho_{XY}=\frac{1}{6}$,则 $\mathrm{Cov}(X,Y)=$________,$D(X-3Y+4)=$________.

(3) 设随机变量 $X\sim \mathrm{Exp}(\lambda)$,且 $D(X)=0.25$,则 $\lambda=$________.

2. 设随机变量的分布律为

X	-1	0	1	2
p_k	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$

求 $E(X),E(-3X+1)$及 $E(X^2)$.

3. 一批零件中有 9 件合格品和 3 件废品.安装机器时从这批零件中任取一件,如果取出的废品不再放回去,求在取得合格品以前已取出的废品数的数学期望.

4. 设在某一规定的时间段里,某电气设备用于最大负荷的时间(单位:min)是一个随机变量,其概率密度为

$$f(x)=\begin{cases}\frac{1}{1500^2}x, & 0\leqslant x\leqslant 1500\\ \frac{1}{1500^2}(3000-x), & 1500<x\leqslant 3000.\\ 0, & \text{其他}\end{cases}$$

求 $E(X)$.

5. 由某种机器切割而成的圆盘的直径(单位:cm)是一个随机变量,其概率密度为

$$f(x)=\begin{cases}\frac{3}{11}(4x-x^2), & 1<x<2\\ 0, & \text{其他}\end{cases}.$$

求圆盘面积的数学期望.

6. 设(X,Y)的分布律为

X \ Y	-1	0	1
1	0.2	0.1	0.1
2	0.1	0	0.1
3	0	0.3	0.1

求:(1)$E(X),E(Y)$;(2)$E\left(\frac{Y}{X}\right)$;(3)$E[(X-Y)^2]$.

7. 设(X,Y)的概率密度为

$$f(x,y)=\begin{cases}12y^2, & 0\leqslant y\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}.$$

求：(1)$E(X)$；(2)$E(Y)$；(3)$E(XY)$.

8. 设随机变量 X 的分布律为

X	-2	0	2
p_k	0.4	0.3	0.3

求：$E(X)$，$D(X)$，$D(\sqrt{10}X-5)$.

9. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\pi\sqrt{1-x^2}}, & |x|<1\\ 0, & |x|\geqslant 1\end{cases}.$$

求：$E(X)$，$D(X)$及 $P\{|X-E(X)|<\sqrt{D(X)}\}$.

10. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}ax, & 0<x<2\\ cx+b, & 2\leqslant x<4.\\ 0, & \text{其他}\end{cases}$$

已知 $E(X)=2$，$P\{1<X<3\}=\dfrac{3}{4}$. 求：(1) a,b,c 的值；(2) $Y=e^X$ 的期望与方差.

11. 某产品的验收方案是从该批产品中任取 6 只产品，若次品数小于等于 1，则该产品通过验收；否则不予通过. 若某工厂该产品的次品率为 0.1，试求在 10 次抽样验收中能通过验收的次数的数学期望.

12. 已知(X,Y)的联合分布律为

X \ Y	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

(1) 试求 $E(X)$，$E(Y)$和 $E(XY)$；(2) 问：X，Y 是否相关？是否独立？

13. 设随机变量(X,Y)服从单位圆上的均匀分布. 验证：X 和 Y 不相关，并且 X 和 Y 也不独立.

14. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{8}(x+y), & 0\leqslant x\leqslant 2, 0\leqslant y\leqslant 2\\ 0, & \text{其他}\end{cases}.$$

求：$E(X)$，$E(Y)$，$D(X)$，$D(Y)$，$\mathrm{Cov}(X,Y)$，ρ_{XY}.

补充与提高

15. 设随机变量 $X_1, X_2, \cdots, X_n (n > 1)$ 独立同分布，且其方差 $\sigma^2 > 0$，令 $Y = \frac{1}{n}\sum_{i=1}^{n} X_i$，则下列正确的是(　　).

(A) $\mathrm{Cov}(X_1, Y) = \frac{\sigma^2}{n}$　　(B) $\mathrm{Cov}(X_1, Y) = \sigma^2$

(C) $D(X_1 + Y) = \frac{n+2}{n}\sigma^2$　　(D) $D(X_1 - Y) = \frac{n+1}{n}\sigma^2$

16. 设二维离散型随机变量 X, Y 的概率分布为

Y / X	0	1	2
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{3}$	0
2	$\frac{1}{12}$	0	$\frac{1}{12}$

(1) 求 $P\{X = 2Y\}$；

(2) 求 $\mathrm{Cov}(X - Y, Y)$.

17. 某厂产品的寿命 T(单位：年)服从指数分布，其概率密度为

$$f(t) = \begin{cases} \frac{1}{4}e^{-\frac{t}{4}}, & t > 0 \\ 0, & t \leqslant 0 \end{cases}.$$

工厂规定，售出的产品在一年内损坏可以调换. 若工厂售出一个产品，能获利 100 元；调换一个产品，工厂要花费 300 元. 试求工厂售出一件产品净获利的数学期望值.

18. 某流水生产线上每个产品不合格的概率为 $p(0 < p < 1)$，各产品合格与否相互独立，当出现一个不合格产品时即停机检修. 设开机后第一次停机时已生产了的产品个数为 X. 求 $E(X)$, $D(X)$.

19. 市场上对某种产品每年的需求量为 X(单位：t)，它服从[2000,4000]上的均匀分布. 已知每出售 1 t 产品可获利 3 万元；售不出去，则每吨需付仓库保管费 1 万元. 试问：每年应生产该产品多少吨，才能使平均收益最大？并求最大平均收益.

20. 在一次拍卖中，两人竞买一幅名画，拍卖以暗标形式进行，并以最高价成交. 设两人的出价相互独立且均服从[1,2]上的均匀分布，求这幅画的期望成交价.

21. 设随机变量 X 的概率密度为 $f(x) = \begin{cases} \frac{1}{2}\cos\frac{x}{2}, & 0 \leqslant x \leqslant \pi \\ 0, & \text{其他} \end{cases}$，对 X 独立重复观察 4 次，Y 表示观察值大于 $\frac{\pi}{3}$ 的次数. 求 Y^2 的数学期望.

22. 某种商品每周内的需求量 X 服从[10,30]上的均匀分布. 已知商店每销售一件该

产品可获利 500 元,若供大于求则削价处理,每处理一件亏损 100 元;若供不应求则缺货部门从别处调剂供应,此时售出一件只获利 300 元. 为使商店平均利润不少于 9280 元,试确定最少进货量.

23. 某车间从一电力公司每天得到的电能 X(单位: kW)服从[10,30]上的均匀分布. 该车间每天对电能的需要量 Y 服从[10,20]上的均匀分布,X 与 Y 独立. 设车间从电力公司得到的每千瓦电能可获得 300 元利润;如车间用电量超过电力公司所提供的数量,就要使用自备发电机提供的附加电能来补充,而从附加电能中每千瓦只能赚取到 100 元利润. 问: 一天中该车间获得利润的数学期望是多少?

24. 设(X,Y)服从二维正态分布,且 $X\sim N(0,3)$,$Y\sim N(0,4)$,$\rho_{XY}=-\dfrac{1}{4}$. 试求 X 和 Y 的联合概率密度.

25. (1) 设 X 为随机变量,c 是任意常数,证明: $D(X)\leqslant E[(X-c)^2]$.

(2) 设 X 是在$[a,b]$上取值的任一随机变量,利用(1)的结论证明 $D(X)\leqslant\left(\dfrac{b-a}{2}\right)^2$.

第 5 章

大数定律与中心极限定理

我们知道,概率统计是研究随机现象统计规律性的学科,而随机现象的统计规律性只有在对随机现象进行大量的考察中才能显现出来. 在第 1 章中已经指出,在相同的条件下,进行大量重复试验时,事件 A 发生的频率随着试验次数的逐渐增大而具有稳定性,但这只是一种直观的说法,并没有从理论上给予证明. 大数定律正是从理论上证明了这一结论的正确性,中心极限定理则从理论上证明了在客观世界中所遇到的许多随机变量往往是近似地服从正态分布的.

5.1 大数定律

在引入大数定律之前,下面首先证明一个重要的不等式.

5.1.1 切比雪夫不等式

定理 5.1(切比雪夫(Chebyshev)不等式) 设随机变量 X 具有数学期望 $E(X)=\mu$,方差 $D(X)=\sigma^2$,则对于任意正数 ε,恒有

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\frac{\sigma^2}{\varepsilon^2}. \tag{5.1}$$

证明 这里仅对 X 是连续型随机变量的情况进行证明,离散型的证明与之类似. 设 X 的概率密度为 $f(x)$,则

$$P\{|X-\mu|\geqslant\varepsilon\}=\int_{|x-\mu|\geqslant\varepsilon}f(x)\mathrm{d}x\leqslant\int_{|x-\mu|\geqslant\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)\mathrm{d}x$$
$$\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2f(x)\mathrm{d}x=\frac{1}{\varepsilon^2}\cdot\sigma^2$$

切比雪夫不等式也可写成如下等价的形式:

$$P\{|X-\mu|<\varepsilon\}\geqslant1-\frac{\sigma^2}{\varepsilon^2}. \tag{5.2}$$

切比雪夫不等式说明,事件$\{|X-E(X)|<\varepsilon\}$发生的概率与方差 $D(X)$关系密切,方差 $D(X)$越小,则 $P\{|X-E(X)|<\varepsilon\}$越大,即随机变量 X 取的值基本上集中在它的期望 $E(X)$附近,这进一步说明了方差的意义.

由于切比雪夫不等式只利用随机变量 X 的期望 $E(X)$及方差 $D(X)$就可以对 X 的概率分布进行估计,因此它在理论研究及实际应用中很有价值.

例 1 已知随机变量 X 的数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$. 当 $\varepsilon=2\sigma$ 和 $\varepsilon=3\sigma$ 时，用切比雪夫不等式求 $P\{|X-\mu|<\varepsilon\}$ 的值至少是多少.

解 根据切比雪夫不等式(5.2)，当 $\varepsilon=2\sigma$ 和 $\varepsilon=3\sigma$ 时，分别有

$$P\{|X-\mu|<2\sigma\}\geqslant 1-\frac{\sigma^2}{(2\sigma)^2}=\frac{3}{4}=0.75,$$

$$P\{|X-\mu|<3\sigma\}\geqslant 1-\frac{\sigma^2}{(3\sigma)^2}=\frac{8}{9}\approx 0.8889.$$

从例 1 可以看出，当随机变量 X 的分布未知时，利用它的数学期望和方差可以知道 $P\{|X-\mu|<\varepsilon\}$ 的值至少是多少，从而可以粗略地估算某些事件发生的概率.

5.1.2 大数定律

定义 5.1 设 $X_1,X_2,\cdots,X_n,\cdots$ 是一随机变量序列，a 是一个常数，若对于任意的正数 ε，有

$$\lim_{n\to\infty}P\{|X_n-a|<\varepsilon\}=1, \tag{5.3}$$

则称随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$ **依概率收敛**于 a，记作 $X_n\xrightarrow{P}a$.

定理 5.2(伯努利大数定律) 设 n_A 是 n 重伯努利试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对于任意正数 ε，有

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|<\varepsilon\right\}=1, \tag{5.4}$$

或

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|\geqslant\varepsilon\right\}=0. \tag{5.5}$$

即

$$\frac{n_A}{n}\xrightarrow{P}p.$$

证明 设随机变量 X_k 为

$$X_k=\begin{cases}1, & 第\ k\ 次试验\ A\ 发生\\ 0, & 第\ k\ 次试验\ A\ 不发生\end{cases}\quad (k=1,2,\cdots,n),$$

则 $n_A=X_1+X_2+\cdots+X_n$，其中 $X_1,X_2,\cdots,X_n$ 相互独立，且都服从以 p 为参数的(0-1)分布，从而

$$E(X_k)=p,\quad D(X_k)=p(1-p),\quad k=1,2,\cdots,n,$$

$$E\left(\frac{n_A}{n}\right)=E\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n}\sum_{k=1}^{n}E(X_k)=p,$$

$$D\left(\frac{n_A}{n}\right)=D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k)=\frac{p(1-p)}{n}.$$

由切比雪夫不等式可得

$$0\leqslant P\left\{\left|\frac{n_A}{n}-p\right|\geqslant\varepsilon\right\}\leqslant\frac{D\left(\frac{n_A}{n}\right)}{\varepsilon^2}=\frac{p(1-p)}{n\varepsilon^2}.$$

因此

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|\geqslant\varepsilon\right\}=0,$$

亦即

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|<\varepsilon\right\}=1.$$

伯努利大数定律从理论上证明了：只要重复独立试验的次数 n 足够多，事件 A 发生的频率$\frac{n_A}{n}$与其概率 p 偏差超过任意正数 ε 的概率都趋于 0，即频率“严重偏离”概率将几乎是不可能的. 这就严密地解释清楚了频率稳定到概率的含义. 正是基于这一理论，当 n 很大时，可以用频率作为概率的近似值.

人们在实践中还发现，除了频率具有稳定性以外，大量观察值的平均值也具有稳定性，这就是切比雪夫大数定律. 实际中常用的是它的特殊情况. 即：

定理 5.3（切比雪夫大数定律的特殊情况） 设随机变量 $X_1,X_2,\cdots,X_n,\cdots$相互独立，服从同一分布，并且具有相同的数学期望和方差：$E(X_k)=\mu,D(X_k)=\sigma^2,k=1,2,\cdots$，则对任意正数 ε，恒有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1, \tag{5.6}$$

即

$$\bar{X}=\frac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{P}\mu.$$

证明 因为

$$E\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n}\sum_{k=1}^{n}E(X_k)=\frac{1}{n}\sum_{k=1}^{n}\mu=\mu,$$

以及

$$D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k)=\frac{1}{n^2}\sum_{k=1}^{n}\sigma^2=\frac{\sigma^2}{n},$$

于是，由切比雪夫不等式，对于任意 $\varepsilon>0$，有

$$0\leqslant P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|\geqslant\varepsilon\right\}\leqslant\frac{D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}.$$

从而

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|\geqslant\varepsilon\right\}=0,$$

即

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1,$$

从而定理得证.

定理 5.3 中要求方差存在，实际上这一条可以去掉，相应的结论仍然成立.

定理 5.4（辛钦大数定律） 设随机变量 $X_1,X_2,\cdots,X_n,\cdots$相互独立，服从同一分布，且

都具有数学期望 $E(X_k)=\mu,k=1,2,\cdots$，则对于任意正数 ε，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1. \tag{5.7}$$

证明略.

辛钦大数定律在实际应用中是很重要的. 辛钦大数定律表明，独立同分布的随机变量的算术平均值 $\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$ 一定稳定于它的期望 $E(\overline{X})=\mu$，这为估计期望值提供了一条切实可行的途径. 如我们要测量某一物理量 μ，在一定的条件下重复测量 n 次得到的测量值 $x_1,x_2,\cdots,x_n$ 是不完全相同的，这些结果可看作是 n 个独立随机变量 $X_1,X_2,\cdots,X_n$ 的实验数值. 显然 $X_1,X_2,\cdots,X_n$ 具有相同的分布，并且有数学期望 μ. 由大数定律可知，当 n 充分大时，n 次测量结果的平均值可作为 μ 的近似值，即

$$\mu\approx\frac{1}{n}\sum_{k=1}^{n}X_k,$$

由此所发生的误差是很小的.

5.2 中心极限定理

在实际问题中，常常需要考虑许多随机因素所产生的总的影响。例如，炮弹射击的落点与目标的偏差，就受着许多随机因素的影响，如瞄准时的误差、空气的阻力所产生的误差、炮弹或炮身结构所引起的误差等。对我们来说重要的是这些随机因素的总的影响。观察表明，如果一个结果是由大量相互独立的随机因素的影响所造成，而每一个别因素在总的影响中所起的作用不大，则这种量一般都服从或近似服从正态分布。本节将讨论在一定的条件下，相互独立的随机变量的和近似服从正态分布的问题，有关的定理称为中心极限定理。这里仅给出两个最常用的中心极限定理。

定理 5.5（独立同分布的中心极限定理） 设随机变量 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，服从同一分布，且都具有数学期望和方差：$E(X_k)=\mu,D(X_k)=\sigma^2>0,k=1,2,\cdots$. 则当 $n\to\infty$ 时，它们的和的极限分布是正态分布，即随机变量之和 $\sum_{k=1}^{n}X_k$ 的标准化变量

$$Y_n=\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}$$

的分布函数 $F_n(x)$ 对于任意 x 满足

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt=\Phi(x). \tag{5.8}$$

证明略.

这就是说均值为 μ、方差为 $\sigma^2>0$ 的独立同分布的随机变量 $X_1,X_2,\cdots,X_n$ 的和 $\sum_{k=1}^{n}X_k$ 的标准化变量，当 n 充分大时，有

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \overset{\text{近似地}}{\sim} N(0,1), \tag{5.9}$$

$$\sum_{k=1}^{n} X_k \overset{\text{近似地}}{\sim} N(n\mu, n\sigma^2). \tag{5.10}$$

将式(5.9)左端改写成 $\dfrac{\frac{1}{n}\sum_{k=1}^{n} X_k - \mu}{\sigma/\sqrt{n}} = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$，这样，上述结果又可写成：当 n 充分大时，

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \overset{\text{近似地}}{\sim} N(0,1) \quad \text{或} \quad \overline{X} \overset{\text{近似地}}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right). \tag{5.11}$$

这是独立同分布中心极限定理结果的另一种形式.这一结果是数理统计中大样本统计推断的基础.

在实际工作中，只要 n 足够大，便可把独立同分布的随机变量之和当作正态变量.

把独立同分布的中心极限定理应用到二项分布上，我们可得到另一个重要的中心极限定理，它是人们最早认识的中心极限定理.

定理 5.6(棣莫弗-拉普拉斯(De Moivre-Laplace)定理)

设随机变量 $\eta_n(n=1,2,\cdots)$服从参数为 $n,p(0<p<1)$的二项分布，则对于任意 x，有

$$\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leqslant x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \Phi(x). \tag{5.12}$$

证明 由于服从二项分布的随机变量 η_n 可视为 n 个相互独立服从同一参数 p 的(0-1)分布的随机变量 $X_1,X_2,\cdots,X_n$ 的和，即有 $\eta_n = \sum_{k=1}^{n} X_k$，其中 $E(X_k) = p, D(X_k) = p(1-p)$，$k=1,2,\cdots,n$，故由独立同分布中心极限定理可得

$$\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leqslant x\right\} = \lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - np}{\sqrt{np(1-p)}} \leqslant x\right\}$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

这个定理表明，正态分布是二项分布的极限分布.当 n 充分大时，二项分布可用正态分布来近似，因而服从二项分布的随机变量 η_n 的概率计算可以近似地用正态随机变量的概率计算，这比直接按二项分布的公式计算概率要简便得多.下面举几个关于中心极限定理应用的例子.

例 1 计算机在进行加法运算时，对每个加数取整(取为最接近于它的整数)，设所有的取整误差是相互独立的，且它们都在(−0.5,0.5)上服从均匀分布.

(1) 若取 1200 个数相加，误差总和的绝对值超过 15 的概率是多少?

(2) 多少个数加在一起使得误差总和的绝对值小于 10 的概率为 0.95 ?

解 设每个加数的取整误差为 $X_k, k=1,2,\cdots,1200$，则 $X_1,X_2,\cdots,X_n$ 相互独立，且都在(−0.5,0.5)上服从均匀分布，并且有

$$E(X_k) = 0, \quad D(X_k) = \frac{1}{12}, \quad k = 1,2,\cdots,1200.$$

(1) 由独立同分布中心极限定理可知,随机变量 $\dfrac{\sum_{k=1}^{1200} X_k - 1200\times 0}{\sqrt{1200}\times\sqrt{1/12}}$ 近似服从 $N(0,1)$,于是所求概率为

$$P\left\{\left|\sum_{k=1}^{1200} X_k\right| > 15\right\} = 1 - P\left\{-15 \leqslant \sum_{k=1}^{1200} X_k \leqslant 15\right\}$$

$$= 1 - P\left\{\frac{-15}{\sqrt{1200}\times\sqrt{\frac{1}{12}}} \leqslant \frac{\sum_{k=1}^{1200} X_k - 1200\times 0}{\sqrt{1200}\times\sqrt{\frac{1}{12}}} \leqslant \frac{15}{\sqrt{1200}\times\sqrt{\frac{1}{12}}}\right\}$$

$$= 1 - P\left\{-1.5 \leqslant \frac{\sum_{k=1}^{1200} X_k}{10} \leqslant 1.5\right\} \approx 1 - (\Phi(1.5) - \Phi(-1.5))$$

$$= 2(1-\Phi(1.5)) = 2(1-0.9332) = 0.1336.$$

(2) 设有 n 个加数,使得

$$P\left\{\left|\sum_{k=1}^{n} X_k\right| < 10\right\} = 0.95,$$

而

$$P\left\{\left|\sum_{k=1}^{n} X_k\right| < 10\right\} = P\left\{\frac{-10}{\sqrt{n/12}} < \frac{\sum_{k=1}^{n} X_k}{\sqrt{n/12}} < \frac{10}{\sqrt{n/12}}\right\}$$

$$\approx \Phi\left(\frac{10}{\sqrt{n/12}}\right) - \Phi\left(-\frac{10}{\sqrt{n/12}}\right)$$

$$= 2\Phi\left(\frac{10}{\sqrt{n/12}}\right) - 1,$$

于是有

$$2\Phi\left(\frac{10}{\sqrt{n/12}}\right) - 1 = 0.95.$$

即

$$\Phi\left(\frac{10}{\sqrt{n/12}}\right) = 0.975 = \Phi(1.96),$$

亦即 $\frac{10}{\sqrt{n}}\sqrt{12} = 1.96$,解得 $n \approx 312.37$,取 $n = 312$ 即可.

例 2 某保险公司有 10 000 人参加人寿保险,每人每年付 12 元保险费.据以往资料,人群中与这项保险业务有关的死亡率为 0.006,死亡后其家属可向保险公司领得 1000 元.试问:保险公司亏本的概率是多少?

解 设 η_n 表示参加保险的 10 000 人中在一年内死亡的人数,则依题意,$\eta_n \sim B(10\,000, 0.006)$.由棣莫弗-拉普拉斯中心极限定理可知,随机变量 $\dfrac{\eta_n - 10\,000\times 0.006}{\sqrt{10\,000\times 0.006\times 0.994}}$

近似服从 $N(0,1)$. 而保险公司亏本的事件为 $\{1000\eta_n > 120\ 000\}$，即 $\{\eta_n > 120\}$，故所求概率为

$$P\{\eta_n > 120\} = 1 - P\{\eta_n \leqslant 120\} = 1 - P\left\{\frac{\eta_n - 60}{\sqrt{59.64}} \leqslant \frac{120 - 60}{\sqrt{59.64}}\right\}$$
$$\approx 1 - \Phi(7.769) \approx 0,$$

即保险公司几乎不会亏本.

例 3 已知某厂生产一大批电子元件，其中合格品占 1/6，现从中任意选取. 问至少应选取多少个这种元件，才能使得选出的这批元件中合格品的比例与 1/6 的差异不大于 0.01 的概率不小于 0.95.

解 设应选取 n 个，并用随机变量 η_n 表示任取的 n 个元件中合格品的个数，则 $\eta_n \sim B\left(n, \frac{1}{6}\right)$. 当 n 很大时，由棣莫弗-拉普拉斯中心极限定理可知，随机变量 $\dfrac{\eta_n - n\times\frac{1}{6}}{\sqrt{n\times\frac{1}{6}\times\frac{5}{6}}}$ 近似服从 $N(0,1)$. 由题意可知，现在要求 n 的值，使得

$$P\left\{\left|\frac{\eta_n}{n} - \frac{1}{6}\right| \leqslant 0.01\right\} \geqslant 0.95,$$

而

$$P\left\{\left|\frac{\eta_n}{n} - \frac{1}{6}\right| \leqslant 0.01\right\} = P\left\{\left|\frac{\eta_n - \frac{1}{6}n}{n}\right| \leqslant 0.01\right\} = P\left\{\frac{\left|\eta_n - n\times\frac{1}{6}\right|}{\sqrt{n\times\frac{1}{6}\times\frac{5}{6}}} \leqslant \frac{0.01\times\sqrt{n}}{\sqrt{\frac{1}{6}\times\frac{5}{6}}}\right\}$$
$$\approx 2\Phi(0.06\sqrt{n/5}) - 1,$$

于是有

$$2\Phi(0.06\sqrt{n/5}) - 1 \geqslant 0.95.$$

即

$$\Phi(0.06\sqrt{n/5}) \geqslant 0.975 = \Phi(1.96),$$

亦即

$$0.06\sqrt{n/5} \geqslant 1.96,$$

解得

$$n \geqslant 5335.56.$$

故取 $n=5336$，即至少应选取 5336 个这种元件.

小　　结

本章应用切比雪夫不等式，以严密的数学形式论证了频率的稳定性，讨论了随机变量序列 $X_1, X_2, \cdots, X_n \cdots$ 的算术平均值 $\frac{1}{n}\sum_{i=1}^{n} X_i$ 依概率收敛于某个常数的大数定律；给出了大量的随机变量之和 $\sum_{i=1}^{n} X_i$ 近似服从正态分布的中心极限定理.

基本要求:

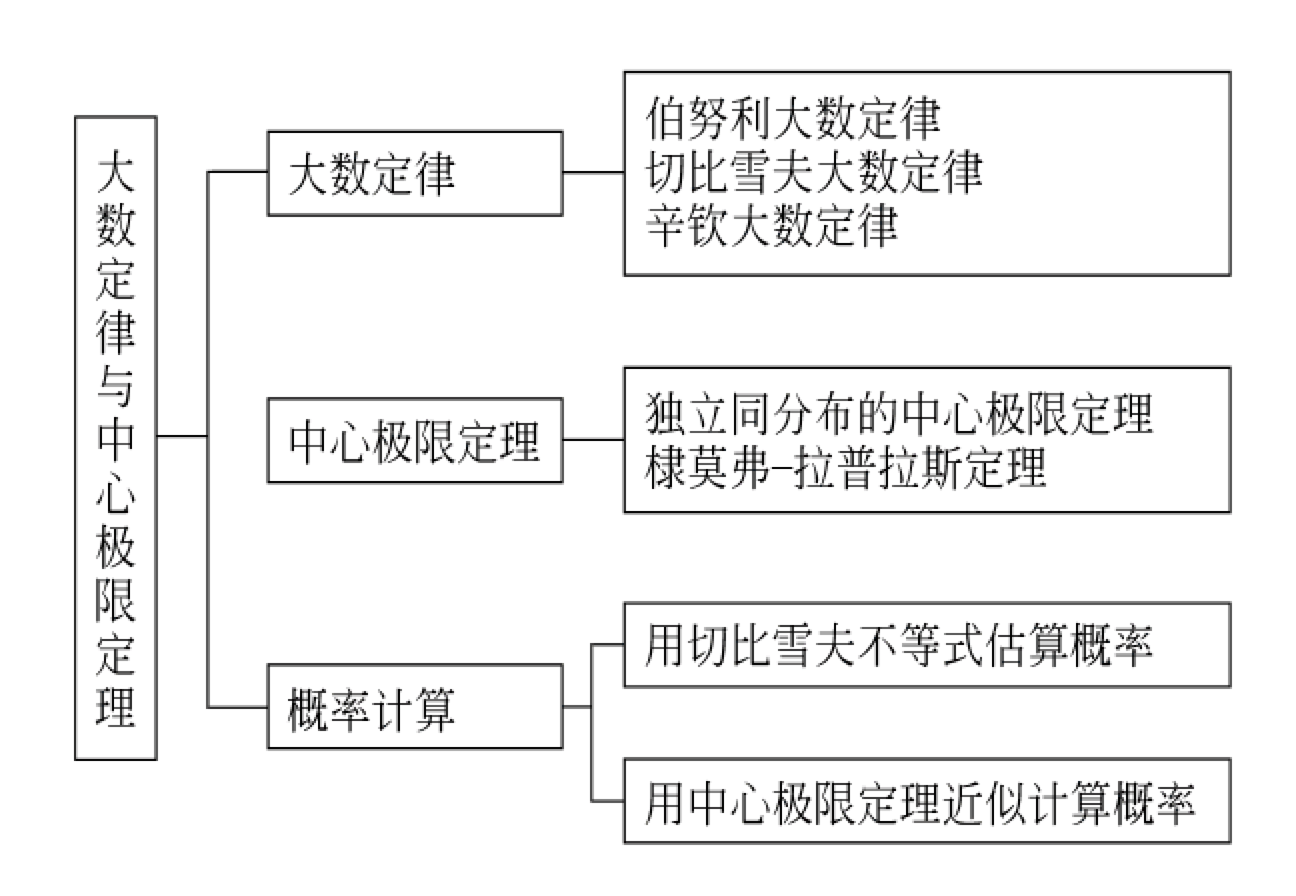
(1) 理解切比雪夫不等式与大数定律的内涵.

(2) 掌握独立同分布的中心极限定理和棣莫弗-拉普拉斯中心极限定理的条件、结论,并熟练应用它们来进行近似计算.

重点与难点:

利用切比雪夫不等式和中心极限定理估计和近似计算一些简单事件的概率.

本章知识结构:



习 题 5

1. 设随机变量 X 的方差为 $D(X)=2$,则根据切比雪夫不等式估计 $P\{|X-E(X)|\geqslant 2\}\leqslant$______.

2. 设 X 为随机变量,且 $E(X)=6, D(X)=4$,则 $P\{0<x<12\}\geqslant$______.

3. 设随机变量 $X_1, X_2, \cdots, X_{100}$ 相互独立同分布,$E(X_k)=1, D(X_k)=16, k=1,2,\cdots,100$. 求 $P\left\{\sum_{k=1}^{100} X_k > 200\right\}$.

4. 设某种电子元件的寿命是相互独立的随机变量,而据以往经验知道,这种电子元件的寿命服从以 $\frac{1}{100}$ 为参数的指数分布. 现从中任意抽取 16 只,求这 16 只元件的寿命总和在 1280~1920 h 之间的概率.

5. 设有 n 个零件,它们的重量 $X_1, X_2, \cdots, X_n$ 都是相互独立、服从同一分布的随机变量,其数学期望为 0.5 kg,均方差为 0.1 kg. 要求 n 只零件的总重量超过 2510 kg 的概率不大于 0.05,则 n 至多是多少?

6. 设某产品的不合格品率为 0.1,现从中任取 100 件,则合格品至少有 85 件的概率是多少?

7. (1) 独立地抛掷一枚均匀的硬币 100 次,出现反面的次数与 50 的差的绝对值不超过 5 的概率是多少?

(2) 要使抛掷一枚均匀的硬币出现反面的频率与 0.5 的差的绝对值不超过 0.005 的概率不小于 0.99,至少需要抛掷多少次?

8. 设电话总机共有 150 个电话分机,每个分机都有 5%的时间使用外线,且是否使用外

线相互独立.要保证每个用户有95%的把握接通外线,总机至少要设置多少条外线?

补充与提高

9. 设 $E(X)=E(Y)=2,D(X)=1,D(Y)=4$,且 $\rho_{XY}=0.5$,则 $P\{|X-Y|\geqslant 6\}\leqslant$______.

10. 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,$S_n=X_1+X_2+\cdots+X_n$,则根据独立同分布中心极限定理,当 n 充分大时,S_n 近似服从正态分布,只要 $X_1,X_2,\cdots,X_n$(　　).

(A) 有相同数学期望　　(B) 有相同方差

(C) 服从同一指数分布　　(D) 服从同一离散分布

11. 为检验一种新药对某种疾病的治愈率为80%是否可靠,给10个患该疾病的病人同时服药,结果治愈人数不超过5人.试判断该药的治愈率为80%是否可靠.

12. 某灯泡厂生产的灯泡的平均寿命为2000 h,标准差为250 h,经技术革新后使平均寿命提高到2250 h,标准差不变.为了确定这一革新成果,任意选取若干只灯泡来作检查,若这些灯泡的平均寿命超过2200 h,就承认革新有效,批准采用革新后的新工艺.问至少应检查多少只灯泡,才能使检查通过的概率超过0.997?

第 6 章

参数估计

从这一章开始，我们将介绍数理统计的内容. 数理统计是一门应用性很强的数学分支，它以概率论的理论为基础，根据试验或观察得到的数据来研究随机现象，对随机现象的客观规律性做出一些合理的估计和判断. 数理统计的核心部分是统计推断，即根据样本包含的信息来建立关于总体的各种结论，它包含两个基本问题：估计问题和假设检验. 这一章我们将主要介绍总体参数的估计问题.

6.1 总体与样本

6.1.1 总体

在数理统计中，通常把被研究对象的某种数量指标的全体称为**总体**. 总体中的每个元素称为**个体**，每个个体是一实数. 例如某工厂生产的灯泡寿命的全体是一总体，每一个灯泡的寿命是一个个体；某高校二年级学生“概率统计”成绩的全体是一总体，每个学生的成绩是一个个体. 对总体的数量指标 X 而言，每个个体所取的值可能是不同的，在试验中，抽取若干个个体就可以观察到 X 的这样或那样的数值，因而这一数量指标 X 是一随机变量. 我们对总体的研究就是对相应随机变量 X 的分布的研究. X 的分布函数和数字特征就称为总体的分布函数和数字特征，今后将不区分总体和相应的随机变量，统称为总体 X.

6.1.2 样本

在实际中，总体的分布一般是未知的. 为了获取对于总体的分布的知识，一般的方法是对总体进行抽样观察(即从总体中抽取一部分个体，逐个观察其数量指标)，从而获得总体 X 的一组数据，借助这组数据采用科学的方法对未知总体进行合理的推断.

从总体 X 中随机地抽取 n 个个体，逐个观察其数量指标，其数量指标为 $X_1, X_2, \cdots, X_n$，这里每个 $X_i (i=1,2,\cdots,n)$ 都是随机变量，称 $X_1, X_2, \cdots, X_n$ 是一个来自总体 X 的**样本**，n 称为这个样本的**容量**. n 次观察一经完成，就得到一组实数 $x_1, x_2, \cdots, x_n$，它们依次是 $X_1, X_2, \cdots, X_n$ 的观察值，称为**样本值**.

抽取样本的目的是为了对总体进行统计推断，因此对抽取的样本有一定的要求. 通常样本需满足两个条件：

(1) $X_1,X_2,\cdots,X_n$ 相互独立;

(2) $X_1,X_2,\cdots,X_n$ 都与总体 X 具有相同的分布.

满足这两个条件的样本 $X_1,X_2,\cdots,X_n$ 称为来自总体 X 的一个**简单随机样本**,简称**样本**.今后所提到的样本都指简单随机样本.

例 1 设 $X_1,X_2,\cdots,X_{16}$ 是总体 $X\sim N(1,4)$ 的样本.求:(1) $Y=\sum_{i=1}^{16}X_i$ 的概率密度;(2) $P\{12<Y<20\}$.

解 由题意知 $X_1,X_2,\cdots,X_{16}$ 独立同服从 $N(1,4)$.

(1) $Y=\sum_{i=1}^{16}X_i$ 服从正态分布 $N(\mu,\sigma^2)$,其中

$$\mu=E(Y)=\sum_{i=1}^{16}E(X_i)=16,$$

$$\sigma^2=D(Y)=\sum_{i=1}^{16}D(X_i)=64.$$

所以

$$Y\sim N(16,64),$$

其密度为

$$f(y)=\frac{1}{\sqrt{2\pi}\cdot 8}e^{-\frac{(y-16)^2}{128}}.$$

(2) $P\{12<Y<20\}=\Phi\left(\frac{20-16}{8}\right)-\Phi\left(\frac{12-16}{8}\right)=\Phi(0.5)-\Phi(-0.5)$
$=2\Phi(0.5)-1=2\times 0.6915-1=0.383.$

6.2 统 计 量

样本是进行统计推断的依据.但在应用时往往不是直接使用样本本身,而是利用样本的适当函数来进行统计推断,为此我们引入统计量的概念.

设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本,$g(X_1,X_2,\cdots,X_n)$ 是样本的连续函数.如果 g 中不包含未知参数,则称 $g(X_1,X_2,\cdots,X_n)$ 是一个**统计量**.若 $x_1,x_2,\cdots,x_n$ 是相应于样本 $X_1,X_2,\cdots,X_n$ 的观察值,则 $g(x_1,x_2,\cdots,x_n)$ 是统计量 $g(X_1,X_2,\cdots,X_n)$ 的观察值.

注:统计量 $g(X_1,X_2,\cdots,X_n)$ 是随机变量的函数,因而是一个随机变量.

下面介绍几种常用的统计量.

设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本,$x_1,x_2,\cdots,x_n$ 是相应的样本值.

(1) **样本均值** 观察值

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\qquad \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i. \tag{6.1}$$

(2) **样本方差** 观察值

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\qquad s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\bar{X}^2\right),\qquad =\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right).\tag{6.2}$$

(3) **样本标准差**　　观察值

$$S=\sqrt{S^2}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2},\qquad s=\sqrt{s^2}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}.\tag{6.3}$$

(4) **样本 k 阶原点矩**　　观察值

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k,\quad k=1,2,\cdots,\qquad a_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k,\quad k=1,2,\cdots.\tag{6.4}$$

例 1　从某厂生产的同种零件中抽得 10 个零件，测得重量（单位：kg）为 10.1，10，9.8，10.5，9.7，10.1，9.9，10.2，10.3，9.9. 求：(1)样本均值 $\bar{x}$；(2)样本方差 s^2.

解

$$\bar{x}=\frac{1}{10}\sum_{i=1}^{10}x_i=10.05;$$

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2=0.0583.$$

6.3　常用统计量的分布

在数理统计中，由总体 X 中获得样本后，通常是借助于样本的函数对未知总体进行统计推断. 为了实现这一目的，需要了解样本的函数所服从的分布，即统计量的分布. 本节介绍几种常用统计量的分布.

6.3.1　χ^2分布

定义 6.1　设 $X_1,X_2,\cdots,X_n$ 是来自总体 $N(0,1)$的样本，则称随机变量

$$\chi^2=X_1^2+X_2^2+\cdots+X_n^2\tag{6.5}$$

服从自由度为 n 的**χ^2分布**，记为 $\chi^2\sim\chi^2(n)$，其中 n 是上式右边独立变量的个数.

可以证明 $\chi^2(n)$分布具有概率密度

$$f_{\chi^2(n)}(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, & x>0\\ 0, & 其他\end{cases},\tag{6.6}$$

其中 $\Gamma(s)=\int_0^{+\infty}e^{-x}x^{s-1}dx\ (s>0)$是 Γ(Gamma 伽马)函数，$f_{\chi^2(n)}(x)$的图形见图 6-1.

可以证明下述结论：

(1) 设 $\chi^2\sim\chi^2(n)$，则 $E(\chi^2)=n,D(\chi^2)=2n$；

(2) $\chi_1^2\sim\chi^2(n_1),\chi_2^2\sim\chi^2(n_2)$，且 χ_1^2,χ_2^2 相互独立，则

$$\chi_1^2+\chi_2^2\sim\chi^2(n_1+n_2).$$

设 $\chi^2\sim\chi^2(n)$，对于给定的正数 $\alpha(0<\alpha<1)$，称满足条件

$$P\{\chi^2>\chi_\alpha^2(n)\}=\alpha$$

的点 $\chi_\alpha^2(n)$ 为 $\chi^2(n)$ 分布的上 α 分位点(见图 6-2).

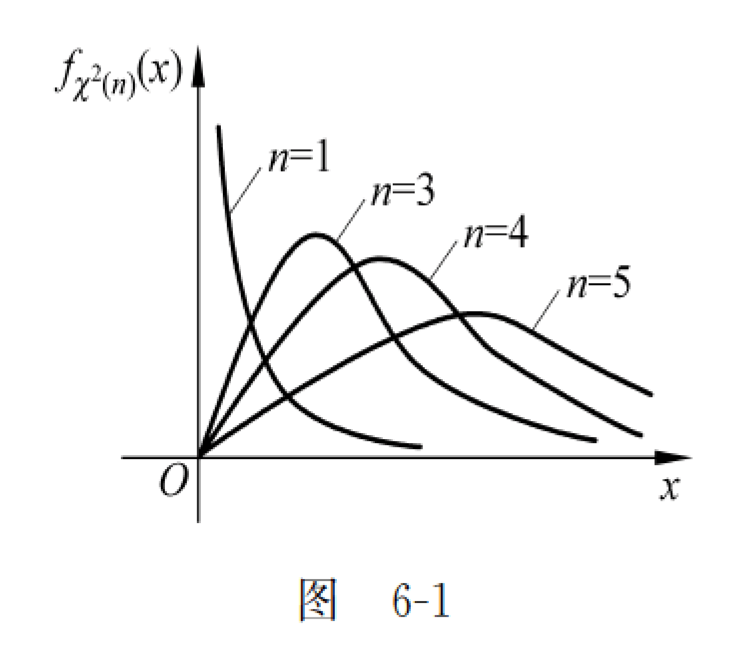


图 6-1

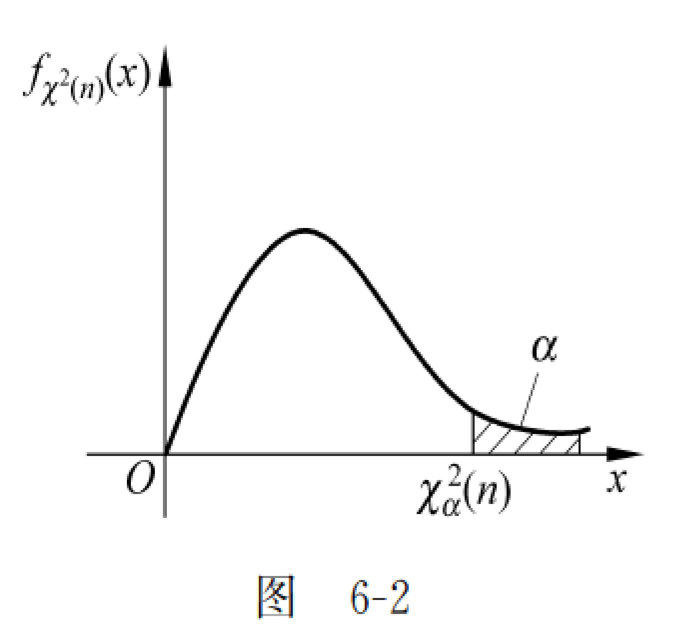


图 6-2

不同的 α,n 的上 α 分位点 $\chi_\alpha^2(n)$ 的值已制成表格可以查用(见附表 A.3). 如 $\chi_{0.01}^2(19)=36.191$.

6.3.2 *t* 分布

定义 6.2 设 $X\sim N(0,1)$, $Y\sim\chi^2(n)$, 且 X,Y 相互独立, 则称随机变量

$$T=\frac{X}{\sqrt{Y/n}} \tag{6.7}$$

服从自由度为 n 的 ***t* 分布**, 记为 $T\sim t(n)$.

t 分布的概率密度为

$$f_T(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\cdot\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},\quad -\infty<x<+\infty. \tag{6.8}$$

$f_T(x)$ 的图形见图 6-3.

可以证明 $\lim\limits_{n\to\infty}f_T(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

设 $T\sim t(n)$, 对于给定的正数 α, $0<\alpha<1$, 称满足条件

$$P\{T>t_\alpha(n)\}=\alpha$$

的点 $t_\alpha(n)$ 为 $t(n)$ 分布的上 α 分位点(见图 6-4).

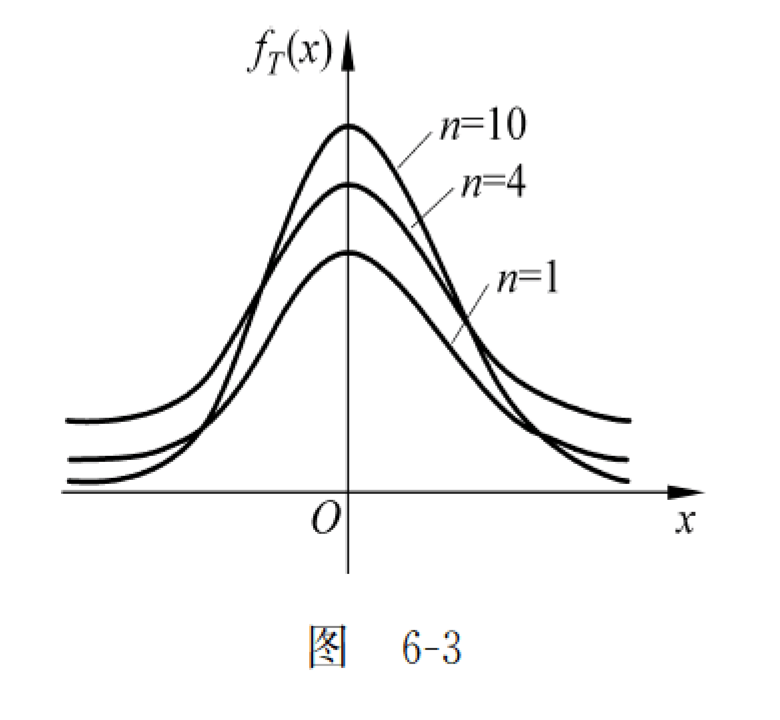


图 6-3

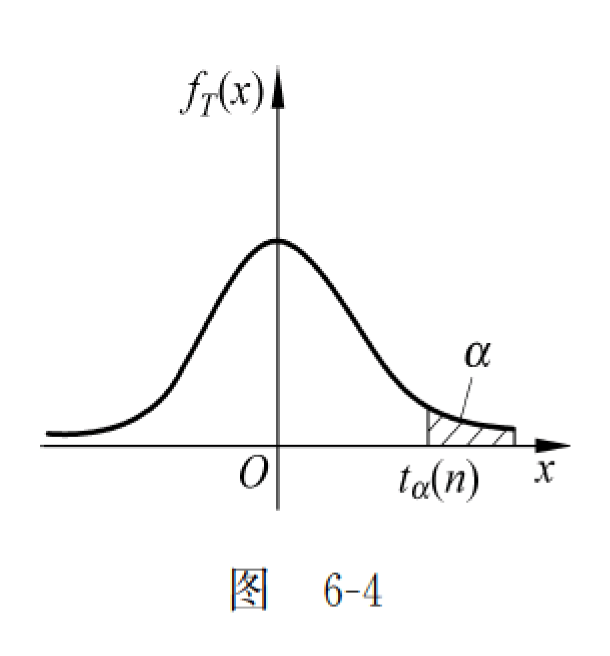


图 6-4

$t_\alpha(n)$ 的值可查附表 A.4, 在 $n>44$ 时可利用 $t_\alpha(n)\approx z_\alpha$ 得到.

6.3.3 *F* 分布

定义 6.3 设 $X\sim\chi^2(m)$, $Y\sim\chi^2(n)$ 且 X,Y 相互独立, 则称随机变量

$$F=\frac{X/m}{Y/n} \tag{6.9}$$

服从自由度为 m,n 的 **F 分布**,记为 $F\sim F(m,n)$.

由定义知道,$F\sim F(m,n)$则$\frac{1}{F}\sim F(n,m)$.

$F(m,n)$分布的概率密度为

$$f_F(x)=\begin{cases}\dfrac{\Gamma\left(\dfrac{m+n}{2}\right)\left(\dfrac{m}{n}\right)^{\frac{m}{2}}x^{\frac{m}{2}-1}}{\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n}{2}\right)\left(1+\dfrac{mx}{n}\right)^{\frac{m+n}{2}}}, & x>0\\ 0, & 其他\end{cases}. \tag{6.10}$$

$f_F(x)$的图形见图 6-5.

设 $F\sim F(m,n)$,对于给定的正数 $\alpha(0<\alpha<1)$,称满足条件 $P\{F>F_\alpha(m,n)\}=\alpha$ 的点 $F_\alpha(m,n)$为 $F(m,n)$分布的上 α 分位点(见图 6-6).

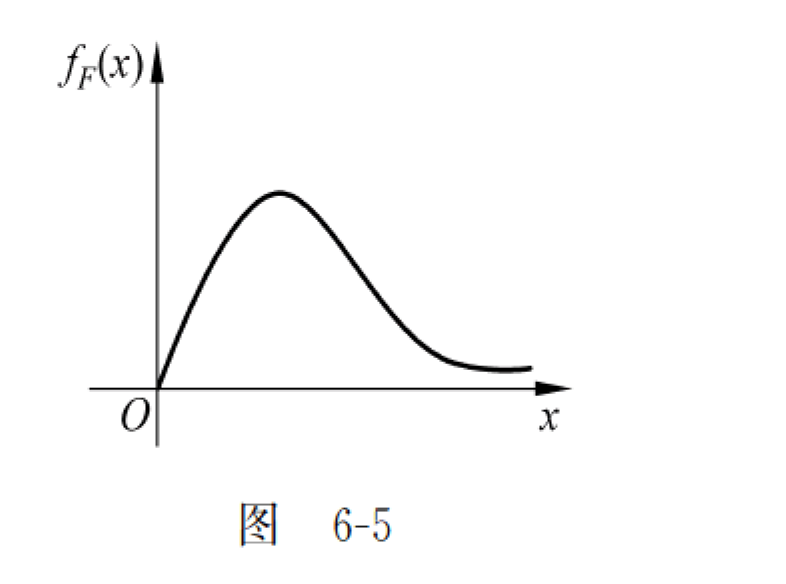


图 6-5

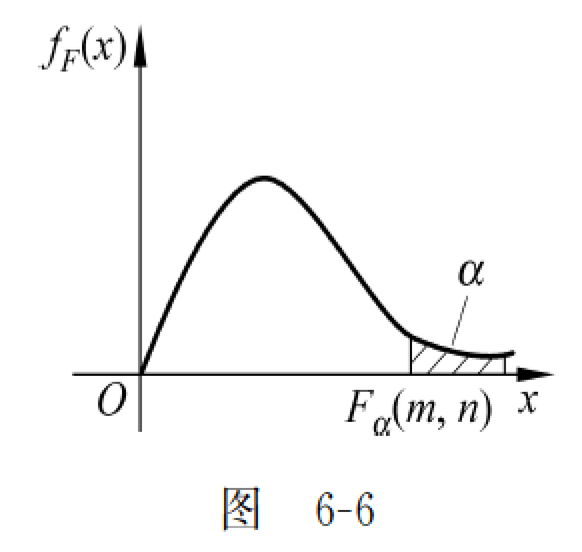


图 6-6

$F_\alpha(m,n)$可由附表 A.5 查到.它具有性质

$$F_{1-\alpha}(m,n)=\frac{1}{F_\alpha(n,m)}.$$

6.3.4 正态总体的统计分布

设 $X_1,X_2,\cdots,X_n$ 是来自正态总体 $N(\mu,\sigma^2)$的样本,则由 3.5 节内容知

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\sim N\left(\mu,\frac{\sigma^2}{n}\right), \tag{6.11}$$

或

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1). \tag{6.12}$$

关于正态总体 $N(\mu,\sigma^2)$的样本均值 $\overline{X}$ 和样本方差 S^2 还有如下结论:

定理 6.1

$$\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1) \tag{6.13}$$

且 $\overline{X}$ 与 S^2 相互独立.(证明见本节附录)

定理 6.2

$$\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1) \tag{6.14}$$

证明　由式(6.12)及式(6.13),按 t 分布定义有

$$\frac{\overline{X}-\mu}{S/\sqrt{n}}=\frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/n-1}}\sim t(n-1).$$

附　　录

定理 6.1 的证明.

记 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$,则有

$$E(\boldsymbol{X})=\begin{bmatrix}\mu\\ \vdots\\ \mu\end{bmatrix},\quad D(\boldsymbol{X})=\sigma^2\boldsymbol{I}.$$

取一个 n 维正交矩阵 $\boldsymbol{A}$,其第一行的每一个元素均为 $1/\sqrt{n}$,如

$$\boldsymbol{A}=\begin{bmatrix}\frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}}\\ \frac{1}{\sqrt{2\times1}} & -\frac{1}{\sqrt{2\times1}} & 0 & \cdots & 0\\ \frac{1}{\sqrt{3\times2}} & \frac{1}{\sqrt{3\times2}} & -\frac{2}{\sqrt{3\times2}} & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & -\frac{n-1}{\sqrt{n(n-1)}}\end{bmatrix}.$$

令 $\boldsymbol{Y}=\boldsymbol{AX}$,则由多维正态分布的性质知 $\boldsymbol{Y}$ 仍服从 n 维正态分布,其均值和方差分别为

$$E(\boldsymbol{Y})=\boldsymbol{A}\cdot E(\boldsymbol{X})=\begin{bmatrix}\sqrt{n}\mu\\ 0\\ \vdots\\ 0\end{bmatrix},$$

$$D(\boldsymbol{Y})=\boldsymbol{A}\cdot\mathrm{var}(\boldsymbol{X})\cdot\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}\cdot\sigma^2\boldsymbol{I}\cdot\boldsymbol{A}^{\mathrm{T}}=\sigma^2\boldsymbol{AA}^{\mathrm{T}}=\sigma^2\boldsymbol{I}.$$

由此,$\boldsymbol{Y}=(Y_1,Y_2,\cdots,Y_n)^{\mathrm{T}}$ 的各个分量相互独立,且都服从正态分布,其方差均为 σ^2,而均值并不完全相同. Y_1 的均值为$\sqrt{n}\mu$, $Y_2,\cdots,Y_n$ 的均值为 0. 注意到 $Y_1=\sqrt{n}\overline{X}$,由于 $\sum_{i=1}^{n}Y_i^2=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\sum_{i=1}^{n}X_i^2$,故而

$$(n-1)S^2=\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}X_i^2-(\sqrt{n}\,\overline{X})^2=\sum_{i=1}^{n}Y_i^2-Y_1^2=\sum_{i=2}^{n}Y_i^2.$$

由于 $Y_2,\cdots,Y_n$ 独立同分布于 $N(0,\sigma^2)$,于是有

$$\frac{(n-1)S^2}{\sigma^2}=\sum_{i=2}^{n}\left(\frac{Y_i}{\sigma}\right)^2\sim\chi^2(n-1).$$

又由于 $\overline{X}=\frac{1}{\sqrt{n}}Y_1$, $S^2=\frac{1}{n-1}\sum_{i=2}^{n}Y_i^2$,所以 $\overline{X}$ 与 S^2 相互独立,这就证明了定理的

式(6.13).

6.4　参数的点估计

参数估计是统计推断的基本问题之一,它包含参数的点估计与参数的区间估计.本节讨论参数的点估计.

6.4.1　参数的点估计的概念

人们常会遇到这样一类问题:已知总体的分布函数,但它的一个或多个参数未知.例如已知总体服从指数分布,但不知其参数 θ 等于多少,这时人们希望借助于总体的一个样本来估计参数 θ.这种问题称为参数的点估计问题.

例 1　某地区去年每月因交通事故死亡的人数如下:

3, 2, 0, 5, 4, 3, 1, 0, 7, 2, 0 ,2

假设每月交通事故死亡的人数 X 服从参数为 λ 的泊松分布,λ 未知,$\lambda>0$.试估计参数 λ.

解　由于 $X\sim P(\lambda)$,所以 $\lambda=E(X)$.

设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim P(\lambda)$的样本,由大数定律知

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}E(X)=\lambda,$$

故可以用 $\overline{X}$ 的观察值 $\bar{x}$ 作为 λ 的估计值.于是 λ 的估计值为 $\bar{x}=\frac{1}{12}(3+2+0+5+4+3+1+0+7+2+0+2)=2.417$.

定义 6.4　设总体 X 的分布已知,但其中有一未知参数 θ,$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本,$x_1,x_2,\cdots,x_n$ 为相应的样本值.构造一个适当的统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$,然后用它的观察值 $\hat{\theta}(x_1,x_2,\cdots,x_n)$来估计 θ 的真值,称 $\hat{\theta}(X_1,X_2,\cdots,X_n)$为 θ 的**估计量**,称 $\hat{\theta}(x_1,x_2,\cdots,x_n)$为 θ的**估计值**,在不致引起混淆的情况下,估计量和估计值简称为**估计**,记为 $\hat{\theta}$.值得注意的是:估计量是样本的函数,估计值是估计量的函数值,对不同的样本值,θ 的估计值往往是不相同的.

6.4.2　点估计的两种常用方法

1. 矩估计法

在例 1 中以样本均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 作为总体均值 $E(X)=\lambda$ 的估计量,即用一阶样本矩作为一阶总体矩的估计量从而求出未知参数 λ 的估计量,这种方法实际上就是矩估计法。

设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本,

$$\mu_k=E(X^k),\quad A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k.$$

由于样本 $X_1,X_2,\cdots,X_n$ 独立同分布，于是 $X_1^k,X_2^k,\cdots,X_n^k$ 也独立同分布，因而

$$E(X_i^k)=E(X^k)=\mu_k,\quad i=1,2,\cdots,n.$$

由大数定律知 $A_k\xrightarrow{P}\mu_k$，于是 n 较大时，$\mu_k\approx A_k$，因此可用 A_k 作为 μ_k 的估计量. 一般地，用样本的 k 阶矩作为总体 k 阶矩的估计量来求未知参数的点估计方法称为**矩估计法**. 所求得的估计量称为**矩估计量**，估计值称为**矩估计值**.

具体做法如下：

设总体 X 的分布已知，其中含有 L 个未知参数 $\theta_1,\theta_2,\cdots,\theta_L$，

$$\mu_k=E(X^k)=\begin{cases}\displaystyle\int_{-\infty}^{+\infty}x^k f(x,\theta_1,\theta_2,\cdots,\theta_L)\mathrm{d}x, & \text{连续总体}\\ \displaystyle\sum_{i=1}^{\infty}x_i^k p(x_i,\theta_1,\theta_2,\cdots,\theta_L), & \text{离散总体}\end{cases}.$$

令 $\mu_k(\theta_1,\theta_2,\cdots,\theta_L)=A_k$，$k=1,2,\cdots,L$，解这个方程组，其解记为

$$\hat{\theta}_k=\hat{\theta}_k(X_1,X_2,\cdots,X_n),\quad k=1,2,\cdots,L,$$

它们可以作为 $\theta_1,\theta_2,\cdots,\theta_L$ 的估计量，即为所求的矩估计。

例 2 已知总体 X 的概率密度为

$$f(x)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}.$$

其中 $\theta>0$，θ 为未知参数，$X_1,X_2,\cdots,X_n$ 为总体 X 的样本，$x_1,x_2,\cdots,x_n$ 为相应的样本值. 求 θ 的矩估计量和矩估计值.

解 $$\mu_1=E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^1 x\sqrt{\theta}x^{\sqrt{\theta}-1}\mathrm{d}x=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}.$$

令

$$\frac{\sqrt{\theta}}{\sqrt{\theta}+1}=A_1=\overline{X},$$

得

$$\theta=\left(\frac{\overline{X}}{1-\overline{X}}\right)^2.$$

所以所求矩估计量为 $\hat{\theta}=\left(\dfrac{\overline{X}}{1-\overline{X}}\right)^2$，矩估计值为 $\hat{\theta}=\left(\dfrac{\overline{x}}{1-\overline{x}}\right)^2$.

例 3 设 X 服从 $[0,\theta]$ 上的均匀分布，θ 未知，$X_1,X_2,\cdots,X_n$ 是总体 X 的一个样本. 求 θ 的矩估计量.

解 因为 X 在 $[0,\theta]$ 上服从均匀分布，所以总体 X 的一阶矩为 $\mu_1=E(X)=\dfrac{\theta}{2}$. 又样本的一阶矩为 $A_1=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i=\overline{X}$，令 $A_1=\mu_1$，得 $\overline{X}=\dfrac{\theta}{2}$. 所以 θ 的矩估计量为 $\hat{\theta}=2\overline{X}$.

例 4 设总体 X 的均值为 μ，方差为 σ^2，其中 μ,σ^2 未知，$X_1,X_2,\cdots,X_n$ 为样本，x_1，

$x_2,\cdots,x_n$ 为样本值. 求 μ,σ^2 的矩估计量.

解 $\mu_1=\mu$, $\mu_2=E(X^2)=D(X)+[E(X)]^2=\sigma^2+\mu^2$.

令

$$\begin{cases}\mu_1=A_1\\ \mu_2=A_2\end{cases},$$

有

$$\begin{cases}\mu=A_1\\ \sigma^2+\mu^2=A_2\end{cases},$$

解得

$$\begin{cases}\mu=A_1=\overline{X}\\ \sigma^2=A_2-\mu^2=A_2-A_1^2\end{cases},$$

所以所求矩估计量为

$$\begin{cases}\hat{\mu}=\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i\\ \hat{\sigma}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2-\overline{X}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\end{cases}.$$

2. 极大似然估计法

极大似然估计法的直观想法：设总体 X 的分布已知，但其中含有未知参数 θ. 如果随机试验的结果得到样本观察值 $x_1,x_2,\cdots,x_n$，则我们选取使这组样本观察值出现的可能性最大的 $\hat{\theta}$ 作为 θ 的估计值.

例 5 已知总体 X 的分布律为 $P\{X=x\}=p^x(1-p)^{1-x}$，$x=0,1$，其中 $0<p<1$ 未知，如果取得样本值为 $x_1=0,x_2=x_3=1$，试估计 p 的值.

解 由题意知，在一次抽样中，事件 $\{X_1=0,X_2=1,X_3=1\}$ 发生了，所以应选使 $P\{X_1=0,X_2=1,X_3=1\}$ 较大的 $\hat{p}$ 为 p 的估计值. 由于

$$\begin{aligned}&P\{X_1=0,X_2=1,X_3=1\}\\=&P\{X_1=0\}\cdot P\{X_2=1\}\cdot P\{X_3=1\}=(1-p)p^2,\end{aligned}$$

当 $p=\dfrac{2}{3}$ 时 $(1-p)p^2$ 最大，所以 p 的估计值为 $\dfrac{2}{3}$.

1) 总体 X 是离散型的情况

设总体 X 是离散型的，它的分布律 $P\{X=x\}=p(x,\theta)$，$\theta\in\Theta$，θ 为待估参数，Θ 为 θ 的取值范围. $X_1,X_2,\cdots,X_n$ 是来自总体的样本，$x_1,x_2,\cdots,x_n$ 是对应的样本值. 事件 $\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}$ 的概率为

$$L(\theta)=L(x_1,x_2,\cdots,x_n,\theta)=\prod_{i=1}^{n}p(x_i,\theta), \tag{6.15}$$

这一概率随 θ 的取值而变化，它是 θ 的函数，$L(\theta)$ 称为样本的**似然函数**. 就固定的样本值 $x_1,x_2,\cdots,x_n$，在 θ 的取值范围内，挑选使概率 $L(\theta)$ 达到最大的参数值 $\hat{\theta}$ 作为 θ 的估计值，即 $L(\hat{\theta})=\max\limits_{\theta\in\Theta}L(\theta)$. 这样的 $\hat{\theta}$ 与样本值 $x_1,x_2,\cdots,x_n$ 有关，记为 $\hat{\theta}(x_1,x_2,\cdots,x_n)$，称为 θ 的**极**

大似然估计值. 相应的统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$称为参数 θ 的**极大似然估计量**.

2) 总体 X 是连续型的情况

设总体 X 是连续型的,其概率密度 $f(x,\theta)$形式已知,且 $\theta\in\Theta$,Θ 是 θ 的取值范围,θ 是待估参数. 设 $X_1,X_2,\cdots,X_n$ 是总体 X 的样本,则样本 $X_1,X_2,\cdots,X_n$ 的联合概率密度为 $\prod\limits_{i=1}^{n} f(x_i,\theta)$,它的大小与$(X_1,X_2,\cdots,X_n)$落在点$(x_1,x_2,\cdots,x_n)$附近的概率大小成正比,当样本值 $x_1,x_2,\cdots,x_n$ 固定时,它是 θ 的函数,记为 $L(\theta)$,并称

$$L(\theta)=L(x_1,x_2,\cdots,x_n,\theta)=\prod_{i=1}^{n} f(x_i,\theta),\quad \theta\in\Theta \tag{6.16}$$

为**似然函数**. 类似于上面的讨论,我们选取使 $L(\theta)$最大的 $\hat{\theta}$ 作为 θ 的估计.

极大似然估计法的一般步骤如下:

(1) 建立似然函数

$$L(\theta)=\prod_{i=1}^{n} p(x_i,\theta) \quad 或 \quad \prod_{i=1}^{n} f(x_i,\theta)$$

(2) 求 $L(\theta)$的极大值点 $\hat{\theta}$,即为所求.

$\left(令\dfrac{\mathrm{d}L(\theta)}{\mathrm{d}\theta}=0 或\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=0 得驻点 \hat{\theta}\right)$

注:有多个待估参数时,可用多元函数求极值法.

例 6 设总体 $X\sim B(10,p)$,$0<p<1$ 是待估参数,$X_1,X_2,\cdots,X_n$ 是来自总体的样本,$x_1,x_2,\cdots,x_n$ 是相应的样本值. 求 p 的极大似然估计值与估计量.

解 因为 $X\sim B(10,p)$,所以 $P\{X=x\}=\dbinom{10}{x}p^x(1-p)^{10-x}$,$x=0,1,\cdots,10$,故似然函数为

$$\begin{aligned} L(p)&=\prod_{i=1}^{n}\binom{10}{x_i}p^{x_i}(1-p)^{10-x_i}\\ &=\prod_{i=1}^{n}\binom{10}{x_i}\cdot p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{10n-\sum\limits_{i=1}^{n}x_i},\end{aligned}$$

$$\ln L(p)=\ln\prod_{i=1}^{n}\binom{10}{x_i}+\left(\sum_{i=1}^{n}x_i\right)\ln p+\left(10n-\sum_{i=1}^{n}x_i\right)\ln(1-p).$$

令

$$\frac{\mathrm{d}\ln L(p)}{\mathrm{d}p}=\frac{\sum\limits_{i=1}^{n}x_i}{p}-\frac{10n-\sum\limits_{i=1}^{n}x_i}{1-p}=0,$$

解得 p 的极大似然估计值为

$$\hat{p}=\frac{1}{10n}\sum_{i=1}^{n}x_i=\frac{\bar{x}}{10},$$

估计量为

$$\hat{p}=\frac{\bar{X}}{10}.$$

例7 设总体 X 的概率密度为

$$f(x,\theta)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases},$$

其中 $\theta>0$ 未知，$x_1,x_2,\cdots,x_n$ 为样本值. 求 θ 的极大似然估计值.

解 (1) 似然函数

$$L(\theta)=\prod_{i=1}^{n}\sqrt{\theta}x_i^{\sqrt{\theta}-1}=\theta^{\frac{n}{2}}\left(\prod_{i=1}^{n}x_i\right)^{\sqrt{\theta}-1}.$$

(2) $\ln L(\theta)=\dfrac{n}{2}\ln\theta+(\sqrt{\theta}-1)\ln\left(\prod_{i=1}^{n}x_i\right).$

令

$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=\frac{n}{2}\cdot\frac{1}{\theta}+\frac{1}{2\sqrt{\theta}}\ln\left(\prod_{i=1}^{n}x_i\right)=0,$$

得极大似然估计值为

$$\hat{\theta}=\left(\frac{-n}{\sum_{i=1}^{n}\ln x_i}\right)^2.$$

例8 设总体 $X\sim N(\mu,\sigma^2)$，μ,σ^2 为未知参数，$x_1,x_2,\cdots,x_n$ 是来自总体 X 的样本值，求 μ,σ^2 的极大似然估计值.

解 X 的概率密度为

$$f(x,\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

$$L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}\mathrm{e}^{\frac{-\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}},$$

$$\ln L(\mu,\sigma^2)=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}.$$

令

$$\begin{cases}\dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\mu}=\dfrac{2\sum_{i=1}^{n}(x_i-\mu)}{2\sigma^2}=0\\[2ex] \dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\sigma^2}=-\dfrac{n}{2}\cdot\dfrac{1}{\sigma^2}+\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4}=0\end{cases},$$

得

$$\begin{cases}\hat{\mu}=\dfrac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}\\[2ex] \hat{\sigma}^2=\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\end{cases}.$$

例 9 设总体 X 的概率密度为 $f(x,\theta)=\begin{cases}e^{-(x-\theta)}, & x\geqslant\theta\\ 0, & 其他\end{cases}$，$\theta$ 未知，$x_1,x_2,\cdots,x_n$ 为样本值. 求 θ 的极大似然估计值.

解 似然函数

$$L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\begin{cases}e^{-\sum\limits_{i=1}^{n}x_i+n\theta}, & x_i\geqslant\theta,\ i=1,2,\cdots,n,\\ 0, & 其他.\end{cases}$$

$$L'(\theta)=ne^{-\sum\limits_{i=1}^{n}x_i+n\theta}>0,\quad x_i\geqslant\theta,\ i=1,2,\cdots,n.$$

$L(\theta)$是 θ 的增函数，在 θ 的取值范围 $\theta\leqslant x_i$ 内，θ 取最大值 $\min\{x_1,x_2,\cdots,x_n\}$时 $L(\theta)$最大，故 θ 的极大似然估计值为 $\hat{\theta}=\min\{x_1,x_2,\cdots,x_n\}$.

6.4.3 估计量的评选标准

由例 2、例 7 可见，对于同一参数，用不同的估计方法求出的估计量可能不相同，我们自然会问，采用哪一个估计量为好呢？因此有必要建立评价估计量好坏的标准. 估计量是一个随机变量，对于不同的样本值，一般会给出参数的不同估计值，因而评价估计量的好坏时，应从某种整体性质去衡量. 下面介绍几个常用的标准.

1. 无偏性

我们希望估计量 $\hat{\theta}$ 的取值不要偏高也不要偏低，对不同样本值，它的值应在 θ 真值附近摆动，即 $\hat{\theta}$ 的平均取值与 θ 的真值基本一致，于是导出了无偏性标准.

定义 6.5 设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$为参数 θ 的估计量，若 $E(\hat{\theta})=\theta$，则称 $\hat{\theta}$ 是 θ 的**无偏估计量**.

例 10 设总体 X 的均值为 μ，方差为 σ^2，$X_1,X_2,\cdots,X_n$ 是来自总体的样本. 证明：

(1) 样本均值 $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ 是 μ 的无偏估计量.

(2) 样本方差 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ 是 σ^2 的无偏估计量.

(3) 矩估计量 $\hat{\sigma}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ 不是 σ^2 的无偏估计量.

证明

(1) 因为

$$E(X_i)=E(X)=\mu,\quad i=1,2,\cdots,n,$$

$$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu,$$

故 $\hat{\mu}=\overline{X}$ 是 μ 的无偏估计量.

(2) $D(X_i)=D(X)=\sigma^2$，$E(X_i^2)=D(X_i)+[E(X_i)]^2=\sigma^2+\mu^2$；

$$D(\overline{X})=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{\sigma^2}{n},E(\overline{X})^2=D(\overline{X})+[E(\overline{X})]^2=\frac{\sigma^2}{n}+\mu^2.$$

于是

$$E(S^2)=E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]=\frac{1}{n-1}E\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)$$
$$=\frac{1}{n-1}\left[n(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]$$
$$=\frac{1}{n-1}\cdot(n-1)\sigma^2=\sigma^2.$$

所以 S^2 是 σ^2 的无偏估计量.

(3) $E(\hat{\sigma}^2)=E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]=E\left(\frac{n-1}{n}S^2\right)=\frac{n-1}{n}\sigma^2\neq\sigma^2$,

所以矩估计量 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 不是 σ^2 的无偏估计量.

2. 有效性

有时一个参数存在许多无偏估计量,选用哪一个好呢?显然应该看它们哪一个取值更集中在真值 θ 附近,即 $\hat{\theta}$ 与 $\theta=E(\hat{\theta})$ 偏离越小越好. 而方差 $D(\hat{\theta})$ 是描述 $\hat{\theta}$ 与 $\theta=E(\hat{\theta})$ 偏离程度的,于是我们有第二个评选标准.

定义 6.6 设 $\hat{\theta}_1,\hat{\theta}_2$ 都是参数 θ 的无偏估计,若 $D(\hat{\theta}_1)<D(\hat{\theta}_2)$,则称 $\hat{\theta}_1$ 较 $\hat{\theta}_2$ **有效**.

例 11 设总体 X 的期望为 μ,方差为 $\sigma^2>0$,X_1,X_2 是样本. 证明:

(1) $\hat{\mu}_1=\frac{1}{3}X_1+\frac{2}{3}X_2$,$\hat{\mu}_2=\overline{X}=\frac{1}{2}(X_1+X_2)$都是 μ 的无偏估计;

(2) $\hat{\mu}_2$ 较 $\hat{\mu}_1$ 有效.

证明

(1) $E(\hat{\mu}_1)=E\left(\frac{1}{3}X_1+\frac{2}{3}X_2\right)=\frac{1}{3}E(X_1)+\frac{2}{3}E(X_2)=\frac{1}{3}\mu+\frac{2}{3}\mu=\mu$,

同理

$$E(\hat{\mu}_2)=\mu,$$

所以 $\hat{\mu}_1,\hat{\mu}_2$ 都是 μ 的无偏估计.

(2) $D(\hat{\mu}_1)=D\left(\frac{1}{3}X_1+\frac{2}{3}X_2\right)=\frac{1}{9}D(X_1)+\frac{4}{9}D(X_2)=\frac{5}{9}\sigma^2$,

$D(\hat{\mu}_2)=\frac{1}{2}\sigma^2$.

由于 $D(\hat{\mu}_2)<D(\hat{\mu}_1)$,所以 $\hat{\mu}_2$ 较 $\hat{\mu}_1$ 有效.

3. 一致性

上述评定估计量好坏的标准——无偏性、有效性,一般都是在所取样本容量固定时提出的. 当样本容量 n 增大时,我们自然希望估计量能充分接近于待估参数,于是有下述评选标准。

定义 6.7 设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$为参数 θ 的估计量,若对任意 $\varepsilon>0$,有

$$\lim_{n\to\infty}P\{|\hat{\theta}-\theta|<\varepsilon\}=1,$$

则称 $\hat{\theta}$ 为 θ 的**一致估计量**.

例如：设 $X_1,X_2,\cdots,X_n$ 是总体 X 的样本，且 $\mu_k=E(X^k)$，$D(X^k)$存在($k=1,2,\cdots,m$)．由大数定律及上述定义知道，样本均值 $\overline{X}$ 是总体值 μ 的一致估计，样本 k 阶矩 A_k 是总体 k 阶矩 μ_k 的一致估计($k=1,2,\cdots,m$)．

6.5 区间估计

前面讨论了参数的点估计，即如何根据样本观察值求得参数的估计值．在实际问题中，不仅需要求出这些参数的估计值，而且往往还需大致估计这些估计值的精确性与可靠性，因此需要引入另一类估计——区间估计．在区间估计理论中，被广泛接受的一种观点是置信区间，它是由奈曼(Neymann)提出的．

6.5.1 置信区间的概念

定义 6.8 设 θ 是总体分布中的未知参数，$\theta\in\Theta$，$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本．对给定的值 $1-\alpha(0<\alpha<1)$，若存在两个统计量 $\hat{\theta}_1(X_1,X_2,\cdots,X_n)$与 $\hat{\theta}_2(X_1,X_2,\cdots,X_n)$，使

$$P\{\hat{\theta}_1<\theta<\hat{\theta}_2\}=1-\alpha,$$

则称随机区间$(\hat{\theta}_1,\hat{\theta}_2)$为 θ 的置信度为 $1-\alpha$ 的**置信区间**，又分别称 $\hat{\theta}_1$ 与 $\hat{\theta}_2$ 为 θ 的**置信下限**与**置信上限**，$1-\alpha$ 称为**置信度**．一旦有了样本值 $x_1,x_2,\cdots,x_n$，区间$(\hat{\theta}_1,\hat{\theta}_2)$也随之确定，它是一个普通区间，也简称为置信区间．

置信区间的意义：若多次抽样(每次抽样容量相同)，则每次抽样都会得到一个置信区间$(\hat{\theta}_1,\hat{\theta}_2)$，这些置信区间，有的包含 θ 的真值，有的不含 θ 的真值，包含 θ 真值的约占有 $100(1-\alpha)\%$．

6.5.2 寻找置信区间的方法

先看一个例子．

例 1 设总体 $X\sim N(\mu,\sigma^2)$，σ^2 已知，μ 未知，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本．求 μ 的置信度为 $1-\alpha$ 的置信区间．

样本均值 $\overline{X}$ 是 μ 的无偏估计，从 $\overline{X}$ 出发选取一个样本与未知参数的函数 $Z=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$(枢轴量)，且有

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1).$$

对事先给定的 $1-\alpha$，确定 a,b，使

$$P\left\{a<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<b\right\}=1-\alpha,$$

显然满足条件的 a,b 不止一对．我们选取 $a=-z_{\alpha/2}$，$b=z_{\alpha/2}$，其中 $z_{\alpha/2}$是标准正态分布上$\dfrac{\alpha}{2}$分位点．由

$$P\left\{-z_{\alpha/2}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_{\alpha/2}\right\}=1-\alpha,$$

有

$$P\left\{\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2}<\mu<\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right\}=1-\alpha.$$

于是

$$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right) \tag{6.17}$$

为 μ 的置信度为 $1-\alpha$ 的置信区间.

若令 $1-\alpha=95\%$,则 μ 的 95%的置信区间为

$$\left(\overline{X}-1.96\frac{\sigma}{\sqrt{n}},\overline{X}+1.96\frac{\sigma}{\sqrt{n}}\right).$$

用该区间估计 μ,我们不仅直接可知估计成功的概率是 95%,且能够以 95%的把握断言:以 $\overline{X}$ 代替 μ 的绝对误差小于 $1.96\frac{\sigma}{\sqrt{n}}$.

若取 $a=-z_{2\alpha/3}$,$b=z_{\alpha/3}$,则又有

$$P\left\{-z_{2\alpha/3}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_{\alpha/3}\right\}=1-\alpha,$$

即

$$P\left\{\overline{X}-\frac{\sigma}{\sqrt{n}}z_{2\alpha/3}<\mu<\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/3}\right\}=1-\alpha.$$

μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/3},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{2\alpha/3}\right), \tag{6.18}$$

而式(6.17)的区间长度为

$$\frac{2\sigma}{\sqrt{n}}z_{\alpha/2},$$

式(6.18)的区间长度为

$$\frac{\sigma}{\sqrt{n}}(z_{\alpha/3}+z_{2\alpha/3}).$$

对不同的 α,我们会发现式(6.17)的置信区间长度较小,表明估计的精确度高.所以取式(6.17)为 μ的置信区间较好.

由例 1 可以看出,寻找参数 θ 的置信区间的具体做法如下:

(1) 选取 θ 的一个较优的估计量 $\hat{\theta}$;

(2) 围绕 $\hat{\theta}$ 寻找一个依赖于样本与 θ 的函数 $Z=Z(X_1,X_2,\cdots,X_n,\theta)$(称为枢轴量),且服从已知分布;

(3) 对给定的置信度 $1-\alpha$,确定 a,b,使 $P\{a<Z<b\}=1-\alpha$,一般选取使 $P\{Z<a\}=P\{Z>b\}=\frac{\alpha}{2}$的 a,b;

(4) 利用不等式变形导出含有 θ 的置信区间$(\hat{\theta}_1,\hat{\theta}_2)$.

6.5.3 正态总体均值与方差的置信区间

1. 单个总体 $N(\mu,\sigma^2)$ 的情况

设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X\sim N(\mu,\sigma^2)$的样本,$\overline{X},S^2$ 分别是样本均值与样本方差,给定置信度为 $1-\alpha$.

1) 均值 μ 的置信区间

(1) σ^2 为已知时

由例 1 得 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right),\quad \text{简记为}\left(\overline{X}\pm\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right). \tag{6.19}$$

(2) σ^2 为未知时

取 $T=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$,由式(6.14)知 $T\sim t(n-1)$. 由

$$P\left\{-t_{\alpha/2}(n-1)<\frac{\overline{X}-\mu}{S/\sqrt{n}}<t_{\alpha/2}(n-1)\right\}=1-\alpha(\text{见图 6-7}),$$

得

$$P\left\{\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)<\mu<\overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right\}=1-\alpha.$$

于是 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}\pm\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right). \tag{6.20}$$

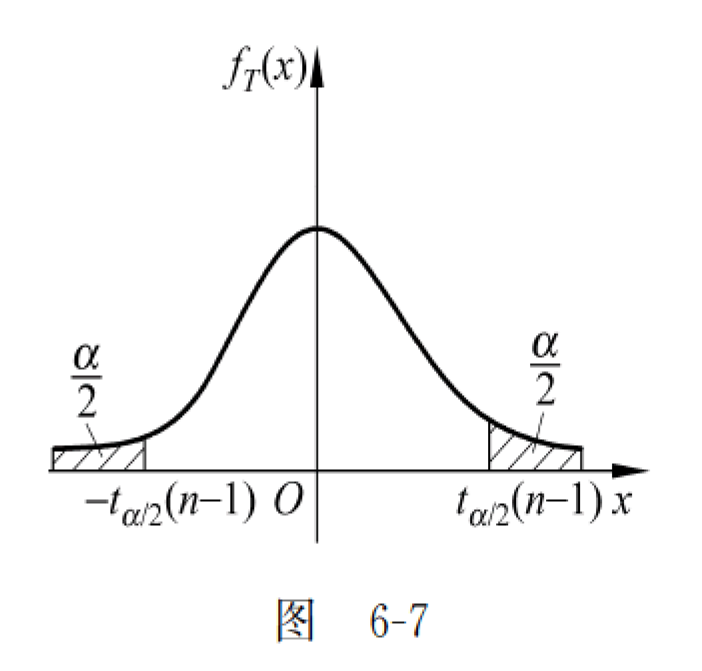


图 6-7

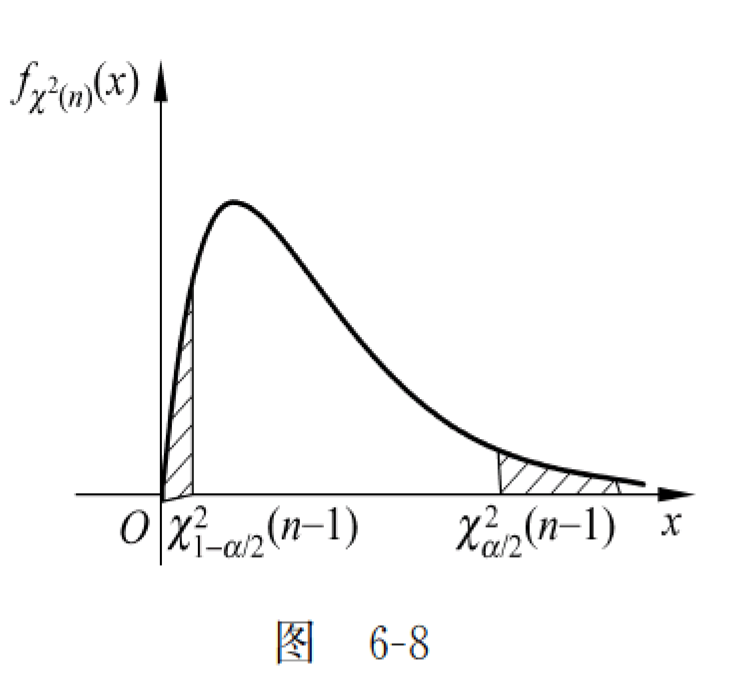


图 6-8

例 2 从一批零件中抽取 9 个零件,测得其直径(单位:mm)分别为 19.7,20.1,19.8,19.9,20.2,20.0,19.9,20.2,20.3,设零件直径服从正态分布 $N(\mu,\sigma^2)$.

在下列条件下,求这批零件直径均值 μ 的置信度为 0.95 的置信区间.

(1) 已知 $\sigma=0.21$;

(2) σ 未知.

解 由题有

$$\bar{x}=20.01,\quad s=0.203,\quad n=9,\quad 1-\alpha=0.95,\quad \alpha=0.05.$$

查表得

$$z_{\alpha/2}=z_{0.025}=1.96,\quad t_{\alpha/2}(8)=t_{0.025}(8)=2.31.$$

(1) μ 的置信度为 $1-\alpha$ 的置信区间为$\left(\overline{X}\pm\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$,代入数据得

$$\left(20.01\pm\frac{0.21}{\sqrt{9}}\times 1.96\right)=(19.87,20.15);$$

(2) μ 的置信区间为$\left(\overline{X}\pm\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$,代入数据得

$$\left(20.01\pm\frac{0.203}{\sqrt{9}}\times 2.31\right)=(19.85,20.17).$$

2) 方差 σ^2 的置信区间

在实际问题中 μ 往往是未知的,所以这里只介绍 μ 未知时 σ^2 的置信区间.

σ^2 的无偏估计为 S^2,由式(6.13)知

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1),$$

故有(见图 6-8)

$$P\left\{\chi^2_{1-\alpha/2}(n-1)<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha/2}(n-1)\right\}=1-\alpha,$$

即

$$P\left\{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}<\sigma^2<\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right\}=1-\alpha,$$

σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right),\tag{6.21}$$

还可得标准差 σ 的置信区间为

$$\left(\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{\alpha/2}(n-1)}},\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right).\tag{6.22}$$

例 3 在例 2 中,求零件直径的方差 σ^2 的置信度为 0.95 的置信区间.

解 由题知

$$\alpha=0.05,\quad n=9,\quad s^2=0.041,$$

查表得

$$\chi^2_{0.025}(8)=17.535,\quad \chi^2_{0.975}(8)=2.18.$$

于是所求置信区间为

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right)=\left(\frac{0.3288}{17.535},\frac{0.3288}{2.18}\right)=(0.0188,0.1508).$$

2. 两个正态总体 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ 的情况

在实际中常常遇到需考虑两个正态总体均值差或方差比的估计问题,为此讨论如下.

定理 6.3 设 $X_1,X_2,\cdots,X_{n_1}$ 是来自总体 $X\sim N(\mu_1,\sigma_1^2)$ 的样本,$Y_1,Y_2,\cdots,Y_{n_2}$ 是来自总体 $Y\sim N(\mu_2,\sigma_2^2)$ 的样本,且这两个样本相互独立;又记 $\overline{X}$ 和 $\overline{Y}$ 分别为它们的样本均值,样本方差分别为

$$S_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2,\quad S_2^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2.$$

则有如下结论：

(1) $\overline{X}-\overline{Y}\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$, (6.23)

或

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1); \qquad (6.24)$$

(2) 当 $\sigma_1^2=\sigma_2^2=\sigma^2$，但 σ^2 未知，有

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2), \qquad (6.25)$$

式中

$$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$$

证明 (1) 由式(6.11)知

$$\overline{X}\sim N\left(\mu_1,\frac{\sigma_1^2}{n_1}\right),\quad \overline{Y}\sim N\left(\mu_2,\frac{\sigma_2^2}{n_2}\right),$$

且 $\overline{X}$ 和 $\overline{Y}$ 相互独立，故式(6.23)成立.

(2) 当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时，由式(6.13)知

$$\frac{(n_1-1)S_1^2}{\sigma^2}\sim \chi^2(n_1-1),\quad \frac{(n_2-1)S_2^2}{\sigma^2}\sim \chi^2(n_2-1),$$

且相互独立，根据 χ^2 分布的可加性，有

$$\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2}\sim \chi^2(n_1+n_2-2).$$

另一方面，当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时，由式(6.24)有

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim N(0,1).$$

由上述两个式子以及 t 分布的定义得式(6.25)成立.

1) 两个正态总体均值差的区间估计

(1) 已知 σ_1^2,σ_2^2，求 $\mu_1-\mu_2$ 的置信区间.

由式(6.24)，对于给定的置信度 $1-\alpha$，有

$$P\left\{\left|\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right|<z_{\alpha/2}\right\}=1-\alpha,$$

于是 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\overline{Y}-\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}z_{\alpha/2},\overline{X}-\overline{Y}+\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}z_{\alpha/2}\right). \qquad (6.26)$$

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知，求 $\mu_1-\mu_2$ 的置信区间，由式(6.25)，仿上有 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\bar{Y}\pm S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\alpha/2}(n_1+n_2-2)\right). \tag{6.27}$$

例 4 研究两种固体燃料火箭推进器的燃烧率,设两者都服从正态分布,并且已知燃烧率的标准差均近似地为 0.05 cm/s,取样本容量为 $n_1=n_2=20$,得燃烧率的样本均值分别为 $\bar{x}=18$ cm/s,$\bar{y}=24$ cm/s.求两燃烧率总体均值差 $\mu_1-\mu_2$ 的置信度为 0.99 的置信区间.

解 由于

$$\sigma_1=\sigma_2=0.05,\quad n_1=n_2=20,\quad \alpha=0.01,$$
$$z_{\alpha/2}=z_{0.005}=2.58,\quad \bar{x}=18,\quad \bar{y}=24,$$

代入式(6.26)得所求置信区间为(−6.04,−5.96).

例 5 为了比较甲、乙两种品牌灯泡的寿命状况(单位:h),抽取了 10 只甲种灯泡和 8 只乙种灯泡,测得平均寿命为 $\bar{x}=1400$ 和 $\bar{y}=1250$,样本标准差分别为 $s_1=52$ 和 $s_2=64$.设两种灯泡的寿命分别服从 $N(\mu_1,\sigma^2)$ 和 $N(\mu_2,\sigma^2)$,其中 μ_1,μ_2,σ^2 均未知,求这两种灯泡平均寿命之差 $\mu_1-\mu_2$ 的 95%置信区间.

解 本题 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知,由题有

$$n_1=10,\quad n_2=8,\quad \alpha=0.05,\quad s_1=52,\quad s_2=64,$$
$$\bar{x}=1400,\quad \bar{y}=1250,\quad t_{\alpha/2}(16)=2.1198,$$
$$s_w^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}=\frac{9.52^2+7.64^2}{10+8-2}=3313,$$
$$s_w\approx 57.5587.$$

代入式(6.27)得所求置信区间为(92.12,207.88).

2) 两个正态总体方差比$\frac{\sigma_1^2}{\sigma_2^2}$的区间估计

我们仅讨论 μ_1,μ_2 未知的情形.

由于

$$\frac{(n_1-1)S_1^2}{\sigma_1^2}\sim\chi^2(n_1-1),\quad \frac{(n_2-1)S_2^2}{\sigma_2^2}\sim\chi^2(n_2-1),$$

由 F 分布定义知

$$F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}=\frac{\dfrac{(n_1-1)S_1^2}{\sigma_1^2}\Big/(n_1-1)}{\dfrac{(n_2-1)S_2^2}{\sigma_2^2}\Big/(n_2-1)}\sim F(n_1-1,n_2-1).$$

对给定置信度 $1-\alpha$,我们有

$$P\left\{F_{1-\alpha/2}(n_1-1,n_2-1)<\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}<F_{\alpha/2}(n_1-1,n_2-1)\right\}=1-\alpha,$$

即

$$P\left\{\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)}<\frac{\sigma_1^2}{\sigma_2^2}<\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right\}=1-\alpha,$$

因此$\frac{\sigma_1^2}{\sigma_2^2}$的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right). \tag{6.28}$$

例 6　在例 5 中求方差比$\frac{\sigma_1^2}{\sigma_2^2}$的 95%的置信区间.

解　因为

$$\alpha=0.05,\quad n_1=10,\quad n_2=8,\quad s_1=52,\quad s_2=64,$$

$$F_{0.025}(9,7)=4.20,\quad F_{0.975}(9,7)=\frac{1}{F_{0.025}(7,9)}=\frac{1}{4.82}\approx 0.207,$$

代入式(6.28)得$\frac{\sigma_1^2}{\sigma_2^2}$的置信度为 95%的置信区间为(0.157,3.182).

小　结

本章主要介绍了数理统计的基本概念和统计推断中的参数估计,主要内容包括:总体与样本;统计量;χ^2 分布、t 分布、F 分布三种抽样分布;正态总体的抽样分布;参数的点估计概念;矩估计法和极大似然估计法两种点估计方法;点估计量的评价标准;参数的区间估计.

基本要求:

(1) 理解总体、个体、样本和统计量的概念.

(2) 了解 χ^2 分布、t 分布、F 分布的定义,会通过查表计算相关概率和三种分布的上 α 分位点;了解正态总体的常用统计量的分布.

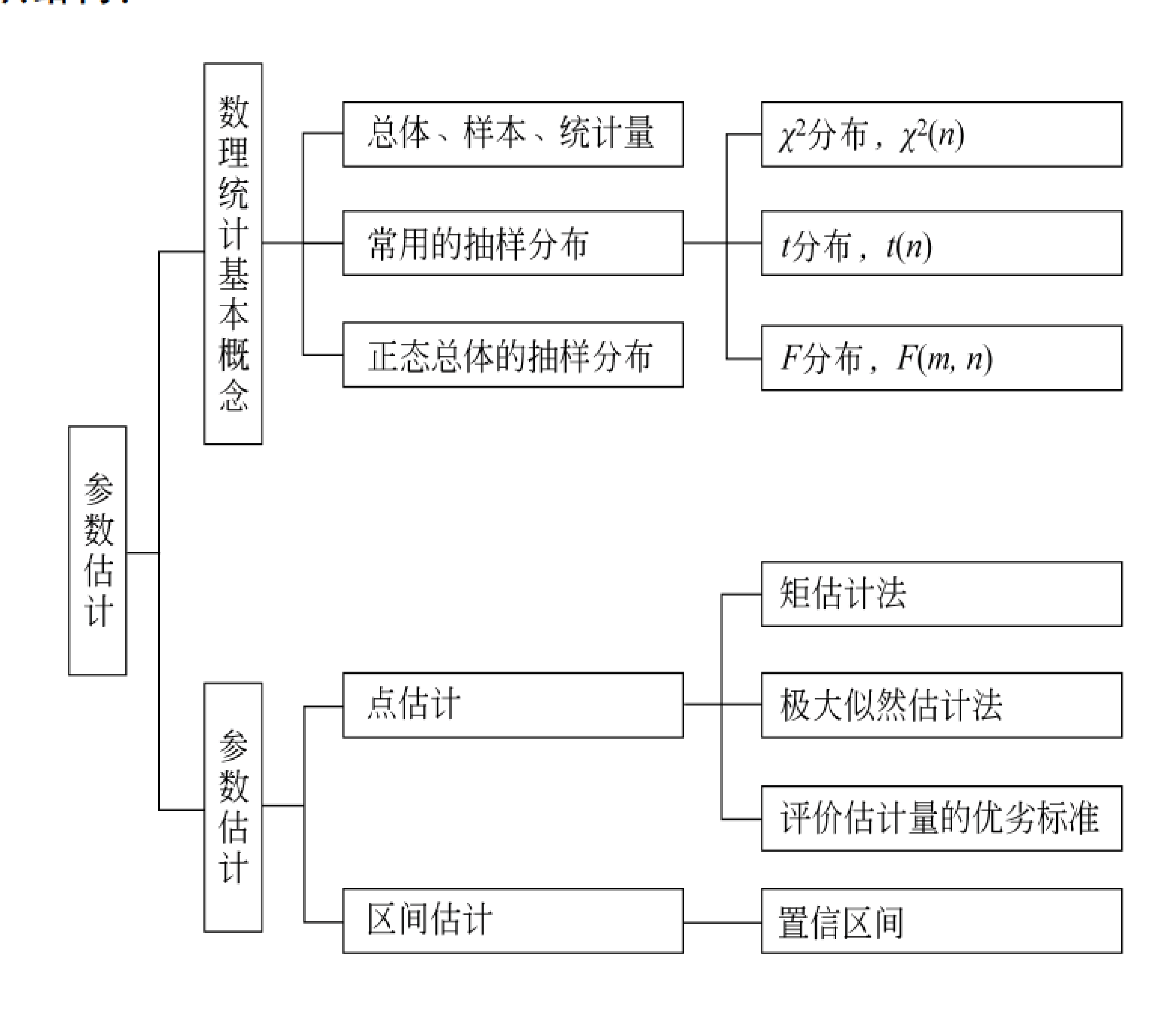
(3) 理解参数点估计的概念;掌握矩估计法与极大似然估计法;了解三种估计量的评选标准.

(4) 理解区间估计的概念;会求正态总体的均值与方差的置信区间.

重点:样本和统计量的概念;矩估计法;极大似然估计法;正态总体均值与方差的区间估计.

难点:极大似然估计法,区间估计的思想和计算.

本章知识结构:



习 题 6

1. 填空题

(1) 设 $X_1,X_2,\cdots,X_n$ 是抽自总体 $X\sim N(\mu,\sigma^2)$ 的样本，$\overline{X},S^2$ 分别为样本均值和样本方差，则 $\overline{X}\sim$______分布；$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim$______分布；$\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim$______分布；$\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)^2\sim$______分布；$\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\overline{X}}{\sigma}\right)^2\sim$______分布.

(2) 设随机变量 $X\sim N(\mu,1)$，$Y\sim\chi^2(n)$，且二者独立，则 $T=\dfrac{X-\mu}{\sqrt{Y/n}}\sim$______分布.

(3) 设 X,Y 相互独立，且 $X\sim\chi^2(8)$，$Y\sim\chi^2(10)$，则 $X+Y\sim$______，$E(X+Y)=$______.

(4) 设 $X_1,X_2,\cdots,X_n$ 为总体 $X\sim B(1,p)$ 的样本，$0<p<1$ 为常数，$\overline{X}$ 为样本均值，则 $P\left\{\overline{X}=\dfrac{k}{n}\right\}=$______.

(5) 已知一批零件的长度 X(单位：cm)服从正态分布 $N(\mu,1)$，从中随机地抽取16个零件，得到长度的平均值为40，则 μ 的置信度为0.95的置信区间是__________.

(由附表知：$\Phi(1.96)=0.975$，$\Phi(1.645)=0.95$)

2. 选择题

(1) 设随机变量 T 服从分布 $t(n)$，对给定的 $\alpha(0<\alpha<1)$，数 $t_\alpha(n)$ 满足 $P\{T>t_\alpha(n)\}=\alpha$. 若 $P\{|T|<t\}=\alpha$，则 t 等于(　　).

(A) $t_{\alpha/2}(n)$　　(B) $t_{1-\alpha/2}(n)$

(C) $t_{\frac{1-\alpha}{2}}(n)$　　(D) $t_{1-\alpha}(n)$

(2) 设一批零件的长度服从 $N(\mu,\sigma^2)$，其中 μ,σ^2 均未知. 现从中随机抽取16个零件，测得样本均值 $\bar{x}=20$ cm，样本标准差 $s=1$ cm，则 μ 置信度为0.90的置信区间是(　　).

(A) $\left(20\pm\dfrac{1}{4}t_{0.05}(16)\right)$　　(B) $\left(20\pm\dfrac{1}{4}t_{0.1}(16)\right)$

(C) $\left(20\pm\dfrac{1}{4}t_{0.05}(15)\right)$　　(D) $\left(20\pm\dfrac{1}{4}t_{0.1}(15)\right)$

3. 设 X_1,X_2,X_3 是来自均值为 μ、方差为 σ^2 的总体的样本. 求：

(1) $D\left(\dfrac{1}{2}X_1+\dfrac{1}{3}X_2+X_3\right)$；

(2) $E(X_1+X_2+X_3)^2$.

4. 设 X_1,X_2,X_3 是总体 $N(\mu,\sigma^2)$ 的样本，其中 μ 已知，σ^2 未知.

(1) 写出 $\dfrac{1}{2}X_1+X_2-\dfrac{1}{3}X_3$ 的概率密度；

(2) 写出 X_1,X_2 的联合概率密度；

(3) 指出 $\dfrac{1}{3}(X_1+X_2+X_3)$，$X_1+X_2-2\mu$，$\max\{X_1,X_2,X_3\}$，$\dfrac{1}{\sigma^2}(X_1^2+X_2^2+X_3^2)$ 中哪些是统计量，哪些不是统计量.

5. 已知样本观察值为15.8，24.2，14.5，17.4，13.2，20.8，19.9，19.1，21.0，18.5，16.4，

22.6,计算样本均值$\overline{X}$及样本方差S^2.

6. 设总体 $X\sim N(40,5^2)$.

(1) 抽取容量为 36 的样本,求样本均值 $\overline{X}$ 落在 38 与 43 之间的概率.

(2) 抽取样本容量多大时,才能使 $P\{|\overline{X}-40|<1\}$达到 0.99?

7. 设 $X_1,X_2,\cdots,X_{16}$是来自总体 $X\sim N(\mu,\sigma^2)$的样本,其中 μ,σ^2 均未知,S^2 为样本方差,求 $P\{S^2/\sigma^2\leqslant 2.041\}$.

8. 设 $X_1,X_2,\cdots,X_{10}$为总体 $X\sim N(0,0.3^2)$的一个样本,求 $P\left\{\sum_{i=1}^{10}X_i^2>1.44\right\}$.

9. 设 $X_1,X_2,\cdots,X_m$ 是来自总体 $X\sim\chi^2(n)$的样本,求样本均值 $\overline{X}$ 的期望及方差.

10. 灯泡厂从某日生产的一批灯泡中抽取 10 个灯泡进行寿命试验,得到灯泡寿命(单位:h)数据如下:

1050,1100,1080,1120,1200,1250,1040,1130,1300,1200.

求该日生产的整批灯泡的平均寿命及寿命方差的无偏估计值.

11. 设 $X_1,X_2,\cdots,X_n$ 为来自总体服从二项分布 $B(m,p)$的样本,m 已知,求 p 的矩估计量和极大似然估计量.

12. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X\sim P(\lambda)$的样本,λ 未知$(\lambda>0)$,求 λ 的矩估计量与极大似然估计量.

13. 设总体 X 的概率密度为

$$f(x,\theta)=\begin{cases}\theta(1-x)^{\theta-1}, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases},$$

其中 θ 未知,$X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本.求 θ 的矩估计量及极大似然估计量.

14. 设总体 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{1}{\theta}x}, & x>0\\ 0, & x\leqslant 0\end{cases},$$

$X_1,X_2,\cdots,X_n$ 为总体 X 的样本.

(1) 求 θ 的矩估计量与极大似然估计量.

(2) 证明所求估计量为 θ 的无偏估计量.

15. 一个电子线路上电压表的读数 X 服从$[\theta,\theta+1]$上均匀分布,其中 θ 是该线路上电压的真值,但它是未知的.假设 $X_1,X_2,\cdots,X_n$ 是此电压表上读数的一组样本.

(1) 证明样本均值 $\overline{X}$ 不是 θ 的无偏估计量.

(2) 求 θ 的矩估计量,证明它是 θ 的无偏估计量.

16. 设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是 θ 的无偏估计量,且 $D(\hat{\theta}_1)=\sigma_1^2$,$D(\hat{\theta}_2)=\sigma_2^2$,取

$$\hat{\theta}=c\hat{\theta}_1+(1-c)\hat{\theta}_2,\quad 0\leqslant c\leqslant 1.$$

(1) 证明 $\hat{\theta}$ 是 θ 的无偏估计量.

(2) 如果 $\hat{\theta}_1,\hat{\theta}_2$ 相互独立,确定 c 使 $D(\hat{\theta})$达到最小.

17. 已知某种清漆的 9 个样品,其干燥时间(单位:h)分别为 6.0,5.7,5.8,6.5,7.0,6.3,5.6,6.1,5.0.设干燥时间 $X\sim N(\mu,\sigma^2)$,求 μ 的置信度为 95%的置信区间.

(1) $\sigma=0.6$ h；

(2) σ 未知.

18. 某工厂生产滚珠，从某日生产的产品中随机抽取 9 个，测得直径的均值为 $\bar{x}=14.91$，直径的标准差 $s=0.203$. 设滚珠直径服从正态分布，求：

(1) 直径的均值 μ 的置信度为 0.95 的置信区间；

(2) 直径的方差 σ^2 的置信度为 0.95 的置信区间.

19. 已知某种材料的抗压强度 $X\sim N(\mu,\sigma^2)$，现随机地抽取 10 个试件进行抗压试验，测得数据如下：482，493，457，471，510，446，435，418，394，469. 求平均抗压强度 μ 的置信水平为 95%的置信区间.

20. 设随机变量 X,Y 相互独立，均服从 $N(0,3^2)$. $X_1,X_2,\cdots,X_9$ 与 $Y_1,Y_2,\cdots,Y_9$ 分别为来自总体 X 和 Y 的样本. 证明：

$$U=\frac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_9^2}}\sim t(9).$$

补充与提高

21. 填空

(1) 设总体 $X\sim N(\mu_1,\sigma^2)$，总体 $Y\sim N(\mu_2,\sigma^2)$. $X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别是来自总体 X 和 Y 的样本，则 $E\left[\dfrac{\sum_{i=1}^{n_1}(X_i-\bar{X})^2+\sum_{j=1}^{n_2}(Y_j-\bar{Y})^2}{n_1+n_2-2}\right]=$______.

(2) 设总体 X 的概率密度为 $f(x)=\dfrac{1}{2}e^{-|x|}$，$-\infty<x<+\infty$，$X_1,X_2,\cdots,X_n$ 为总体 X 的样本，其样本方差为 S^2，则 $E(S^2)=$______.

22. 选择

(1) 设随机变量 X 和 Y 都服从标准正态分布，则(　　).

(A) $X+Y$ 服从正态分布　　(B) X^2+Y^2 服从 χ^2 分布

(C) X^2 和 Y^2 都服从 χ^2 分布　　(D) $\dfrac{X^2}{Y^2}$ 服从 F 分布

(2) 设 $X_1,X_2,\cdots,X_n(n\geqslant 2)$ 为来自总体 $X\sim N(0,1)$ 的简单随机样本，$\bar{X}$ 为样本均值，S^2 为样本方差，则(　　).

(A) $n\bar{X}\sim N(0,1)$　　(B) $nS^2\sim\chi^2(n)$

(C) $\dfrac{(n-1)\bar{X}}{S}\sim t(n-1)$　　(D) $\dfrac{(n-1)X_1^2}{\sum_{i=2}^{n}X_i^2}\sim F(1,n-1)$

23. 设某种元件的使用寿命 X 的概率密度为

$$f(x,\theta)=\begin{cases}2e^{-2(x-\theta)}, & x>\theta\\ 0, & x\leqslant\theta\end{cases},$$

其中 $\theta>0$ 未知. 又设 $x_1,x_2,\cdots,x_n$ 是 X 的一组样本观测值. 求参数 θ 的极大似然估计值.

24. 设总体 X 服从 $N(\mu,\sigma^2)$，从该总体中抽取样本 $X_1,X_2,\cdots,X_{2n}$，$n\geqslant 2$，其样本均值为 $\overline{X}=\dfrac{1}{2n}\sum\limits_{i=1}^{2n}X_i$，求统计量 $Y=\sum\limits_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$ 的数学期望 $E(Y)$.

25. 设总体 X 的概率分布为

X	0	1	2	3
p	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta\left(0<\theta<\dfrac{1}{2}\right)$是未知参数，利用总体 X 的样本值 3，1，3，0，3，1，2，3 求 θ 的矩估计值和极大似然估计值.

26. 设总体 X 的分布函数为

$$F(x,\beta)=\begin{cases}1-\dfrac{1}{x^{\beta}}, & x>1\\ 0, & x\leqslant 1\end{cases},$$

其中未知参数 $\beta>1$，$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本. 求：(1) β 的矩估计量；(2) β 的极大似然估计量.

27. 设总体 X 的概率密度为

$$f(x,\theta)=\begin{cases}\dfrac{\theta^2}{x^3}e^{-\frac{\theta}{x}}, & x>0\\ 0, & \text{其他}\end{cases},$$

其中 θ 为未知参数且 $\theta>0$，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本.

(1) 求 θ 的矩估计量；　　(2) 求 θ 的极大似然估计量.

28. 假设 0.50，1.25，0.80，2.00 是来自总体 X 的样本值. 已知 $Y=\ln X$ 服从正态分布 $N(\mu,1)$. 求：

(1) X 的数学期望 $E(X)$；

(2) μ 的置信度为 0.95 的置信区间；

(3) 利用上述结果求 $b=E(X)$的置信度为 0.95 的置信区间.

29. 欲比较甲、乙两种棉花品种的优劣，现假设用它们纺出的棉纱强度分别服从 $N(\mu_1,2.18^2)$和$N(\mu_2,1.76^2)$. 试验者从这两种棉纱中分别抽取样本 $X_1,X_2,\cdots,X_{200}$ 和 $Y_1,Y_2,\cdots,Y_{100}$，其均值为 $\overline{X}=5.32$，$\overline{Y}=5.76$. 求均值差 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间.

30. 随机地从 A 批导线中抽取 4 根，从 B 批导线中抽取 5 根，测得电阻(单位：Ω)为

A：0.143，0.142，0.143，0.137；

B：0.140，0.142，0.136，0.138，0.140.

设 A，B 两批导线的电阻分别服从正态分布 $N(\mu_1,\sigma_1^2)$和 $N(\mu_2,\sigma_2^2)$，$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 未知. 求：

(1) 当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时，求 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间.

(2) 方差比$\dfrac{\sigma_1^2}{\sigma_2^2}$的置信度为 0.95 的置信区间.

第 7 章

假设检验

7.1 假设检验的基本概念

在第 6 章中我们介绍了统计推断的基本问题之一——参数估计. 本章将介绍另一基本问题——假设检验,即在总体的分布未知或只知其形式,但不知其参数时,对总体的分布或分布参数做出某种假设,然后根据所得样本,运用统计分析的方法来检验这一假设是否正确,从而做出接受或拒绝的决定.

7.1.1 假设检验的基本思想及做法

下面先结合例子来说明假设检验的思想与做法.

例 1 已知在正常生产情况下某种汽车零件的重量 X 服从正态分布 $N(54,0.75^2)$,在某日生产的零件中抽取 10 件,测得重量如下:54.0,55.1,53.8,54.2,52.1,54.2,55.0,55.8,55.1,55.3,如果标准差不变,能否认为该日生产的零件的平均重量为 54?

解 该问题中,已知总体 $X\sim N(\mu,\sigma^2)$ 且 $\sigma=0.75$ 不变,需要确定该日 $\mu=54$ 是否成立. 为此,我们可以提出

$$H_0:\mu=\mu_0=54,\quad H_1:\mu\neq\mu_0=54. \tag{7.1}$$

下面只需由获得的数据,做出接受 H_0 或拒绝 H_0 的决策. 如果接受 H_0 就可以认为该日生产的零件的平均重量为 54,否则不是.

由于要检验的是总体均值 μ,故首先想到能否用样本均值 $\overline{X}$ 来进行判断. 由于总体 $X\sim N(\mu,\sigma^2)$,所以 $Z=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$,当 H_0 为真时 $\mu=\mu_0$,此时 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$(其中 $\sigma=0.75,n=10$)是一个统计量,且有 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$. 由于样本均值 $\overline{X}$ 是总体均值 μ_0 的无偏估计,因此当 H_0 为真时,$\overline{X}$ 的观察值 $\bar{x}$ 与 μ_0 很接近,$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 的观察值应在 0 附近摆动,所以 $|Z|=\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|$ 观察值应较小. 另一方面,若 H_1 为真,即 $\mu\neq\mu_0$,则 $\overline{X}$ 的观察值 $\bar{x}$ 与 μ 很接近,因此 $\bar{x}$ 与 μ_0 的偏差应较大,$|Z|=\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|$ 观察值较大.

综上，H_0 为真时$|Z|=\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|$观察值较大，是不应常出现的，在一次抽样中发生的可能性应较小．记 $A=\left\{|Z|=\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant K\right\}$，则在 H_0 为真时 A 应是一个小概率事件，在一次抽样中，不应该发生．若发生了则有理由怀疑 H_0 为真的真实性，于是应拒绝接受 H_0，接受 H_1；若 A 未发生则没有理由怀疑 H_0 的正确性，于是应接受 H_0．给定 α，$0<\alpha<1$，一般 α 应较小，当 H_0 为真时，按 $P(A)=P\left\{|Z|=\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant K\right\}=\alpha$，可得 $K=z_{\alpha/2}$，于是小概率事件为 $A=\left\{|Z|=\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant z_{\alpha/2}\right\}$．

在本题中，若取 $\alpha=0.05$，$z_{\alpha/2}=z_{0.025}=1.96$，由题得

$$\bar{x}=54.46,\quad \sigma=0.75,\quad n=10,$$

$$|z|=\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|=\left|\frac{54.46-54}{0.75/\sqrt{10}}\right|=1.94<1.96=z_{\alpha/2}.$$

所以在一次抽样中 A 未发生，故接受 H_0，即认为该日生产的零件平均重量为 54．

上述做法的步骤如下：

(1) 由问题本身提出假设

$$H_0:\mu=\mu_0=54,\quad H_1:\mu\neq\mu_0.$$

H_0 称为**原假设**，H_1 为**备择假设**．

(2) 在 H_0 为真时，选择统计量 $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$．

(3) H_0 为真时，确定拒绝 H_0 的形式

$$\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant k.$$

(4) 给定 α（α 较小），由 $P\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant k\right\}=\alpha$ 定出 $k=z_{\alpha/2}$．

(5) 由于 $A=\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geqslant z_{\alpha/2}\right\}$是小概率事件，由实际推断原理来判断做出拒绝或接受 H_0 的决定．数 α 称为**显著性水平**；$Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$称为**检验统计量**．当统计量取某个区域 C 中的值时，我们拒绝 H_0，称区域 C 为**拒绝域**，拒绝域的边界点称为**临界点**．如例 1 中拒绝域 $|Z|\geqslant z_{\alpha/2}$，而 $Z=-z_{\alpha/2}$，$Z=z_{\alpha/2}$为临界点．

7.1.2 双边假设检验与单边假设检验

形如式(7.1)中的备择假设 $H_1:\mu\neq\mu_0$，表示 μ 可能大于 μ_0，也可能小于 μ_0，称为双边备择假设．而形如式(7.1)的假设检验为**双边假设检验**．有时我们只关心总体均值是否增大，例如，试验新工艺以提高材料的强度．这时，所考虑的总体均值应该越大越好．如果我们能判断在新工艺下总体均值较以往正常生产的大，则可考虑采用新工艺．此时，需要检验假设

$$H_0: \mu = \mu_0, \quad H_1: \mu > \mu_0 \tag{7.2}$$

称为**右边检验**.

类似地,检验假设

$$H_0: \mu = \mu_0, \quad H_1: \mu < \mu_0 \tag{7.3}$$

称为**左边检验**.

7.1.3 假设检验可能犯的两类错误

由于在检验中,我们是依据小概率事件在一次试验中几乎不发生的这一原理由样本进行决策的,但几乎不可能发生,不表示一定不发生,所以决策可能犯错误.一种可能是在假设 H_0 实际上为真时,犯拒绝 H_0 的错误,称这类"弃真"的错误为**第Ⅰ类错误**,因为是按小概率事件在一次抽样中是否发生来判定 H_0 是否成立,所以犯第Ⅰ类错误的概率为小概率事件的概率 α;另一种可能是 H_0 实际上不真时,犯接受 H_0 的错误,称这类"取伪"的错误为**第Ⅱ类错误**.

在确定检验法则时,我们应尽可能使犯两类错误的概率都较小.但是,在样本容量固定时,若减小犯一类错误的概率,则犯另一类错误的概率往往增大.若要使犯两类错误的概率都减小,除非增加样本容量.样本容量固定时,一般,我们总是控制犯第Ⅰ类错误的概率 α,这种假设检验问题称为**显著性检验问题**.

7.1.4 参数假设检验的步骤

参数假设检验的步骤如下:

(1) 由实际问题,提出原假设 H_0 及备择假设 H_1;

(2) 确定检验统计量以及所服从的分布和拒绝域的形式;

(3) 在给定显著性水平 α 下,按 $P\{拒绝 H_0 | H_0 为真\}=\alpha$,求出拒绝域;

(4) 由抽样结果计算检验统计量的观察值,进而确定接受 H_0 还是拒绝 H_0.

7.2 正态总体参数的假设检验

7.2.1 单个正态总体参数的假设检验

1. 关于正态总体 $N(\mu, \sigma^2)$ 均值 μ 的假设检验

1) $H_0: \mu=\mu_0, H_1: \mu\neq\mu_0$(双边假设)

(1) σ^2 已知时,见例1,检验统计量为

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1),$$

此时 H_0 的拒绝域为

$$|Z| \geqslant z_{\alpha/2}. \tag{7.4}$$

(2) σ^2 未知时,我们取检验统计量为

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1).$$

H_0 为真时，$|T|$ 的观察值应在 0 附近摆动，H_1 为真时其观察值应偏离 0. 故 H_0 为真时，H_0 拒绝形式 $|T| \geqslant k$，由 $p\{|T| \geqslant k\} = \alpha$ 可确定 $k = t_{\alpha/2}(n-1)$，于是原假设 H_0 的拒绝域为

$$|T| \geqslant t_{\alpha/2}(n-1). \tag{7.5}$$

2) $H_0: \mu = \mu_0, H_1: \mu > \mu_0$（右边假设）

(1) σ^2 已知时，取检验统计量为

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1).$$

当 H_0 为真时，z 的观察值应落在 0 附近；而 H_1 为真时，z 的观察值应偏大. 因此 H_0 的拒绝域形式为 $Z \geqslant k$，由 $p\{拒绝\ H_0 | H_0\ 为真\} = p\{Z \geqslant k\} = \alpha$ 确定得 $k = z_\alpha$，于是拒绝域为

$$Z \geqslant z_\alpha. \tag{7.6}$$

(2) σ^2 未知时，我们取检验统计量为

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1),$$

仿(1)可得 H_0 的拒绝域为

$$T \geqslant t_\alpha(n-1). \tag{7.7}$$

3) $H_0: \mu = \mu_0, H_1: \mu < \mu_0$（左边假设）

仿 2)中的方法有如下结论.

(1) σ^2 已知时，取检验统计量为

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1),$$

可得 H_0 的拒绝域为

$$Z \leqslant -z_\alpha. \tag{7.8}$$

(2) σ^2 未知时，我们取检验统计量为

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1),$$

可得 H_0 的拒绝域为

$$T \leqslant -t_\alpha(n-1). \tag{7.9}$$

现将这些情况列于下表：

σ^2 已知时，关于 μ 的检验（**Z 检验**）结论

原假设 H_0	备择假设 H_1	统计量及其分布	H_0 的拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$	$\lvert Z \rvert \geqslant z_{\alpha/2}$
$\mu = \mu_0$	$\mu > \mu_0$		$Z \geqslant z_\alpha$
$\mu = \mu_0$	$\mu < \mu_0$		$Z \leqslant -z_\alpha$

σ^2 未知时，关于 μ 的检验（**T 检验**）结论

原假设 H_0	备择假设 H_1	统计量及其分布	H_0 的拒绝域
$\mu=\mu_0$	$\mu\neq\mu_0$	$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$	$\lvert T\rvert\geqslant t_{\alpha/2}(n-1)$
$\mu=\mu_0$	$\mu>\mu_0$		$T\geqslant t_\alpha(n-1)$
$\mu=\mu_0$	$\mu<\mu_0$		$T\leqslant -t_\alpha(n-1)$

注：在 Z 检验及 T 检验中，单边检验原假设换成 H_0：$\mu\leqslant\mu_0$（或 $\mu\geqslant\mu_0$）时，对应的 H_0 拒绝域不变.

例 1 某工厂用自动包装机包装葡萄糖，规定标准重量为每袋净重 500 g，现在随机地抽取 10 袋，算得 $\bar{x}=501.3$，$s=5.62$. 设每袋净重服从正态分布，则包装机工作是否正常？（取 $\alpha=0.05$）

解 按题意，需检验 $H_0:\mu=\mu_0=500$，$H_1:\mu\neq\mu_0=500$.

由于 σ^2 未知，所以 H_0 的拒绝域为

$$\mid T\mid=\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right|\geqslant t_{\alpha/2}(n-1),$$

现在 $n=10$，$t_{0.025}(9)=2.26$，$\bar{x}=501.3$，$s=5.62$，所以

$$\mid t\mid=\left|\frac{501.3-500}{5.62/\sqrt{10}}\right|=0.731<2.26,$$

因此接受 H_0，认为包装机工作正常.

例 2 某种电子元件的寿命 X（单位：h）服从正态分布，μ，σ^2 均未知. 观测得 16 只元件的寿命如下：159，280，101，212，214，379，179，264，222，362，168，250，149，260，485，170，是否有理由认为元件的平均寿命大于 225（单位：h）？（取 $\alpha=0.05$）

解 按题意需检验

$$H_0:\mu\leqslant\mu_0=225,\quad H_1:\mu>\mu_0=225.$$

由于 σ^2 未知，所以 H_0 的拒绝域为

$$T=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\geqslant t_\alpha(n-1),$$

现在 $n=16$，$t_{0.05}(15)=1.7531$，$\bar{x}=241.5$，$s=98.7256$，即有

$$t=\frac{241.5-225}{98.7256/\sqrt{16}}=0.6685<1.7531.$$

它不落在拒绝域内，故接受 H_0，即认为元件的平均寿命不大于 225.

2. 关于正态总体 $N(\mu,\sigma^2)$ 的方差 σ^2 的假设检验

设总体 $X\sim N(\mu,\sigma^2)$，μ，σ^2 均未知，$X_1,X_2,\cdots,X_n$ 是来自总体的样本，这里仅讨论双边检验：

$$H_0:\sigma^2=\sigma_0^2,\quad H_1:\sigma^2\neq\sigma_0^2. \tag{7.10}$$

考虑样本方差 s^2 是总体方差 σ^2 的无偏估计，又

$$\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1).$$

当 H_0 为真时，$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$ 是一统计量，且 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$，所以我们取 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$ 为检验统计量. H_0 为真时 S^2 的观察值应在 σ_0^2 附近摆动，即 χ^2 的观察值应落在 $n-1$ 附近；反之 H_1 为真时，χ^2 的观察值应偏离 $n-1$. 因此 H_0 的拒绝域形式应为

$$(\chi^2\leqslant k_1)\cup(\chi^2\geqslant k_2),\quad k_1,k_2\text{ 为常数}.$$

给定显著水平 α，确定 k_1,k_2，使得

$$P\{(\chi^2\leqslant k_1)\cup(\chi^2\geqslant k_2)\}=\alpha.$$

习惯上取

$$P\{\chi^2\leqslant k_1\}=\frac{\alpha}{2},$$

$$P\{\chi^2\geqslant k_2\}=\frac{\alpha}{2}.$$

按 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$，由上式可得(见图 7-1)

$$k_1=\chi^2_{1-\alpha/2}(n-1),\quad k_2=\chi^2_{\alpha/2}(n-1).$$

故得(7.10)中 H_0 的拒绝域为

$$\chi^2\leqslant\chi^2_{1-\alpha/2}(n-1)\quad\text{或}\quad\chi^2\geqslant\chi^2_{\alpha/2}(n-1).$$

同理可得单边检验的情况，在此不再叙述. 现将对 σ^2 的检验列于下表：

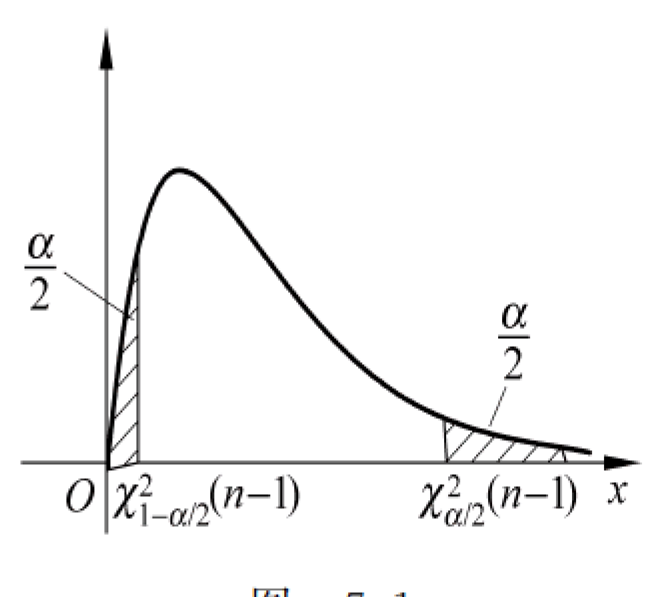


图 7-1


μ 未知时，关于 σ^2 的检验(χ^2 检验)结论

原假设 H_0	备择假设 H_1	统计量及其分布	H_0 的拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\chi^2\geqslant\chi^2_{\alpha/2}(n-1)$ 或 $\chi^2\leqslant\chi^2_{1-\alpha/2}(n-1)$
$\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\geqslant\chi^2_{\alpha}(n-1)$
$\sigma^2=\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\leqslant\chi^2_{1-\alpha}(n-1)$

例 3 在上面的例 1 中，能否认为每袋葡萄糖净重的标准差为 $\sigma=5$？(单位：g)(取 $\alpha=0.05$)

解 检验假设

$$H_0:\sigma=\sigma_0=5,\quad H_1:\sigma\neq\sigma_0=5.$$

由于 μ 未知，H_0 的拒绝域为 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\geqslant\chi^2_{\alpha/2}(n-1)$ 或 $\chi^2\leqslant\chi^2_{1-\alpha/2}(n-1)$，由题有 $\sigma_0=5,n=10,s=5.62,\alpha=0.05,\chi^2_{0.025}(9)=19.0,\chi^2_{0.975}(9)=2.70$，

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}=\frac{9\times5.62^2}{5^2}=11.37,$$

因为 $\chi^2_{0.975}(9)\leqslant\chi^2\leqslant\chi^2_{0.025}(9)$，所以接受 H_0，即可以认为葡萄糖净重的标准差为 5 g.

7.2.2 两个正态总体参数的假设检验

1. 两个正态总体均值的假设检验

设 $X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别为来自正态总体 $N(\mu_1,\sigma_1^2)$ 和 $N(\mu_2,\sigma_2^2)$ 的样本，

样本方差分别为 S_1^2, S_2^2. 我们需检验假设

$$H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 \neq \mu_2.$$

(1) 假设 σ_1^2, σ_2^2 已知，由式(6.24)知

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1),$$

因此，当 H_0 为真时，

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1).$$

H_0 为真时，拒绝 H_0 的形式为 $\frac{|\overline{X}-\overline{Y}|}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \geqslant k$. 给定 α，按 $P\left(\frac{|\overline{X}-\overline{Y}|}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \geqslant k\right)=\alpha$ 得 $k=z_{\alpha/2}$，所以拒绝域为

$$\frac{|\overline{X}-\overline{Y}|}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \geqslant z_{\alpha/2}.$$

(2) 假设 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知，由式(6.25)知

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{1/n_1+1/n_2}} \sim t(n_1+n_2-2),$$

因此当 H_0 为真时，

$$\frac{\overline{X}-\overline{Y}}{S_w\sqrt{1/n_1+1/n_2}} \sim t(n_1+n_2-2).$$

拒绝 H_0 的拒绝域为

$$\frac{\overline{X}-\overline{Y}}{S_w\sqrt{1/n_1+1/n_2}} \geqslant t_{\alpha/2}(n_1+n_2-2).$$

由上述方法，关于两个总体均值的假设检验，有以下结果.

① 已知 σ_1^2, σ_2^2 时，有

原假设 H_0	备择假设 H_1	统计量及其分布	H_0 的拒绝域
$\mu_1=\mu_2$	$\mu_1\neq\mu_2$	$Z=\frac{\overline{X}-\overline{Y}}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$	$\|Z\|\geqslant z_{\alpha/2}$
$\mu_1\leqslant\mu_2$	$\mu_1>\mu_2$		$Z\geqslant z_\alpha$
$\mu_1\geqslant\mu_2$	$\mu_1<\mu_2$		$Z\leqslant -z_\alpha$

② $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知时，有

原假设 H_0	备择假设 H_1	统计量及其分布	H_0 的拒绝域
$\mu_1=\mu_2$	$\mu_1\neq\mu_2$	$T=\frac{\overline{X}-\overline{Y}}{S_w\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2)$	$\|T\|\geqslant t_{\alpha/2}(n_1+n_2-2)$
$\mu_1\leqslant\mu_2$	$\mu_1>\mu_2$		$T\geqslant t_\alpha(n_1+n_2-2)$
$\mu_1\geqslant\mu_2$	$\mu_1<\mu_2$		$T\leqslant -t_\alpha(n_1+n_2-2)$

注：在上述单侧检验中，原假设改为 $H_0: \mu_1=\mu_2$ 时拒绝域不变.

例 4　比较 A, B 两种小麦的蛋白质含量，随机抽取 A 种小麦 10 个样本，测得

$\overline{x}=14.3, s_1^2=1.62$;随机抽取 B 种小麦 5 个样本,测得 $\overline{y}=11.7, s_2^2=0.14$. 假定这两种小麦的蛋白质含量都服从正态分布,且有相同的方差. 试在 $\alpha=0.01$ 水平下,检验两种小麦的蛋白质含量有无差异.

解 按题意需检验假设

$$H_0: \mu_1=\mu_2, \quad H_1: \mu_1 \neq \mu_2.$$

由于 $\sigma_1^2=\sigma_2^2$ 未知,所以 H_0 的拒绝域为

$$|T|=\frac{|\overline{X}-\overline{Y}|}{S_w \sqrt{1/n_1+1/n_2}} \geqslant t_{\alpha/2}(n_1+n_2-2).$$

由题有

$$\overline{x}=14.3, \quad s_1^2=1.62, \quad \overline{y}=11.7, \quad s_2^2=0.14,$$

按

$$s_w^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

得

$$s_w^2=0.2, \quad s_w=0.447, \quad t_{\alpha/2}(n_1+n_2-2)=t_{0.005}(13)=3.0123,$$

故

$$|t|=\left|\frac{14.3-11.7}{0.447\sqrt{1/10+1/5}}\right|=10.62>t_{\alpha/2}(13)=3.0123.$$

因此拒绝 H_0,应当认为两种小麦蛋白质含量有差异.

2. 两个正态总体方差的假设检验

设有两个正态总体 $N(\mu_1,\sigma_1^2)$ 和 $N(\mu_2,\sigma_2^2)$,这里仅介绍 μ_1,μ_2 未知时方差的检验.

检验假设

$$H_0: \sigma_1^2=\sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2. \tag{7.11}$$

设 $X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别为来自正态总体 $N(\mu_1,\sigma_1^2)$ 和 $N(\mu_2,\sigma_2^2)$ 的样本,且它们是相互独立的. 由式(6.13)和 F 分布定义知 $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1,n_2-1)$,当 H_0 为真时,

$$\frac{S_1^2}{S_2^2} \sim F(n_1-1,n_2-1), \tag{7.12}$$

且 $\dfrac{S_1^2}{S_2^2}$ 不应偏大或偏小. 因此,当 H_0 为真时,H_0 的拒绝形式为

$$\frac{S_1^2}{S_2^2} \geqslant k_1 \quad \text{或} \quad \frac{S_1^2}{S_2^2} \leqslant k_2.$$

按

$$P\left(\frac{S_1^2}{S_2^2} \geqslant k_1\right)=P\left(\frac{S_1^2}{S_2^2} \leqslant k_2\right)=\frac{\alpha}{2}$$

有

$$k_1=F_{\alpha/2}(n_1-1,n_2-1), \quad k_2=F_{1-\alpha/2}(n_1-1,n_2-1),$$

故假设式(7.11)中的 H_0 的拒绝域为

$$\frac{S_1^2}{S_2^2} \geqslant F_{\alpha/2}(n_1-1,n_2-1) \quad \text{或} \quad \frac{S_1^2}{S_2^2} \leqslant F_{1-\alpha/2}(n_1-1,n_2-1). \tag{7.13}$$

类似地,我们有假设检验

$$H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2. \tag{7.14}$$

H_0 的拒绝域为

$$\frac{S_1^2}{S_2^2}\geqslant F_\alpha(n_1-1,n_2-1). \tag{7.15}$$

值得注意的是式(7.14)中的原假设换成 $H_0:\sigma_1^2\leqslant\sigma_2^2$ 时,H_0 的拒绝域仍为式(7.15).

例 5 有两台机器生产金属部件,分别在两台机器所生产的部件中各取容量 $n_1=30,n_2=20$ 的样本,测得部件重量的样本方差分别为 $s_1^2=15.46,s_2^2=9.66$. 设两样本相互独立,两总体分别服从 $N(\mu_1,\sigma_1^2),N(\mu_2,\sigma_2^2)$分布. 试在水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma_1^2=\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2.$$

解 由式(7.15)知 H_0 的拒绝域为$\frac{S_1^2}{S_2^2}\geqslant F_\alpha(n_1-1,n_2-1)$,由题意知 $n_1=30,n_2=20,\alpha=0.05$. 查表得 $F_{0.05}(29,19)=2.07$,而$\frac{s_1^2}{s_2^2}=\frac{15.46}{9.66}=1.60<F_{0.05}(29,19)$,所以应接受 H_0.

小 结

本章主要内容包括:假设检验的基本概念与基本原理;正态总体参数的假设检验.

基本要求:

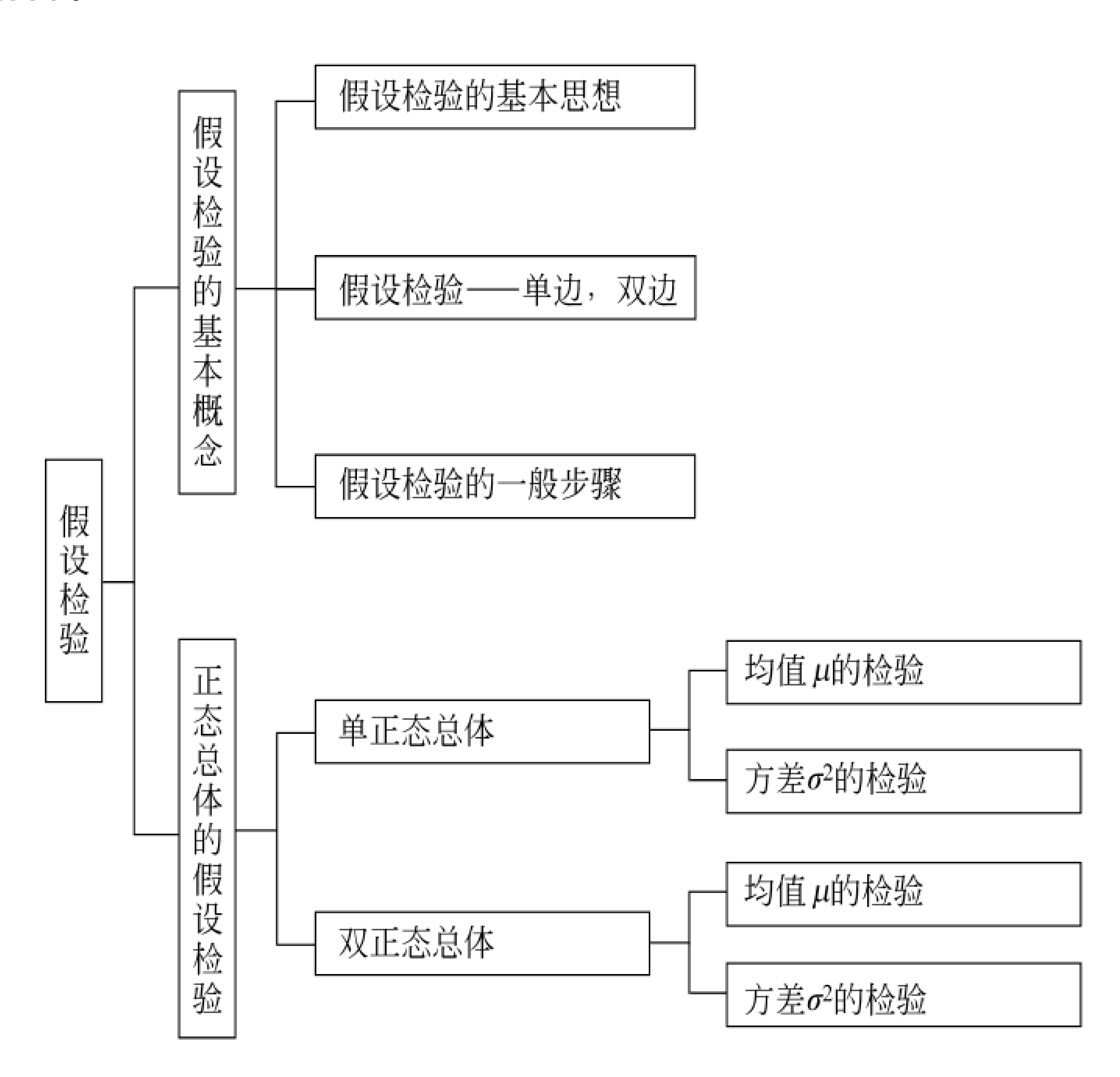
(1) 理解假设检验的基本思想;掌握假设检验的基本步骤;了解假设检验可能产生的两类错误.

(2) 掌握正态总体的均值与方差的假设检验.

重点:假设检验的基本思想;正态总体均值与方差的假设检验.

难点:假设检验的计算推导.

本章知识结构:



习 题 7

1. 某批矿砂的 5 个样本中的镍含量,经测定为(%)

3.25,3.27,3.24,3.26,3.24,

设测定值总体服从正态分布.问在 $\alpha=0.01$ 水平下能否接受假设:这批矿砂的镍含量的均值为 3.25.

2. 进行了 5 次试验,测得锰的熔点(单位:℃)如下:

1269,1271,1256,1265,1254.

已知锰的熔点服从正态分布,在 $\alpha=0.01$ 水平下是否可以认为锰的熔点显著高于 1250℃?

3. 设某次考试的学生成绩服从正态分布,从中随机抽取 30 位考生的成绩,算得平均成绩为 66.5 分,标准差为 15 分.在显著性水平 $\alpha=0.05$ 下是否可以认为这次考试全体考生的平均成绩为 70 分?

4. 假设香烟中尼古丁含量服从正态分布 $N(\mu,\sigma^2)$,现从某牌香烟中随机抽取 20 支,其尼古丁含量的平均值 $\bar{x}=18.6$mg,样本标准差 $s=2.4$mg,在 $\alpha=0.01$ 水平下,以下两种情况下能否接受"该牌的香烟尼古丁含量不超过 17.5 mg"的断言?

(1) $\sigma=2.4$ 已知;

(2) σ 未知.

5. 用过去的铸造方法所造零件的强度均值是 52.8 g/mm,标准差是 1.6 g/mm;为降低成本,改变了铸造方法,随后抽取 9 个样品,测其强度(单位:g/mm)为 51.9,53.0,52.7,54.7,53.2,52.3,52.5,51.1,54.1.假设强度服从正态分布,试判断强度均值是否有改变.(取显著性水平 $\alpha=0.05$)

6. 某种金属丝,根据长期正常生产的累积资料知道其折断力服从正态分布,方差为 64 kg^2.最近从一批产品中抽取 10 根作折断力试验,测得结果(单位:kg)如下:

578,572,570,508,572,570,572,596,584,570.

问:在显著性水平 $\alpha=0.05$ 下,这批金属丝的折断力的方差变化了吗?

7. 某炼铁厂铁水的含碳量 X 在正常情况下服从正态分布.现对操作工艺进行了某些改变,从中抽取 7 炉铁水的试样,测得含碳量数据如下:

4.421,4.052,4.357,4.394,4.326,4.287,4.683.

问是否可以认为新工艺炼出的铁水含碳量的方差仍为 0.112^2.(取显著性水平 $\alpha=0.05$)

8. 在正常情况下,维尼纶纤度服从正态分布,标准差不大于 0.048.某日抽取 5 根纤维,测得纤度为

1.32,1.55,1.36,1.40,1.44.

问:该日维尼纶纤度的标准差是否正常?(取显著性水平 $\alpha=0.01$)

9. 测定某种溶液中的水分,它的 10 个测定值给出 $s=0.037\%$.设测定值总体为正态总体 $N(\mu,\sigma^2)$,试在 $\alpha=0.05$ 水平下检验假设 H_0: $\sigma^2=0.04\%$,H_1:$\sigma^2<0.04\%$.

10. 某公司生产的发动机部件的直径服从正态分布,该公司称它的标准差 $\sigma=0.048$ cm.现随机抽取 5 个部件,测得它们的直径为 1.32,1.55,1.36,1.40,1.44,取显著性水平 $\alpha=0.05$.问:

(1) 我们能否认为该公司生产的发动机部件的直径的标准差确实为 0.048 cm?

(2) 我们能否认为 $\sigma^2 \leqslant 0.048^2$?

补充与提高

11. 考虑正态总体 $N(\mu,\sigma^2)$ 和假设检验 $H_0:\sigma^2=\sigma_0^2, H_1:\sigma^2>\sigma_0^2$. 证明：当 μ 未知时，H_0 的拒绝域为 $\frac{(n-1)S^2}{\sigma_0^2} \geqslant \chi_\alpha^2(n-1)$.

12. 假设有 A, B 两种药，试验者欲比较它们服用 2 小时后血液中的含量是否一样. 对药品 A，随机抽取 8 个病人，他们服药 2 小时后，测得血液中药的浓度(单位：mg/ml)为

1.23, 1.42, 1.41, 1.62, 1.55, 1.51, 1.60, 1.76;

对药品 B，随机抽取 6 个病人，他们服药 2 小时后，测得血液中药的浓度(单位：mg/ml)为

1.76, 1.41, 1.87, 1.49, 1.67, 1.81.

假定这两组观察值服从具有公共方差的正态分布，试在显著性水平 $\alpha=0.10$ 下，检验病人血液中这两种药的浓度是否有显著不同.

13. 甲、乙两厂生产同一种电阻，现从甲、乙两厂的产品中分别随机抽取 12 个和 10 个样品，测得它们的电阻值后，计算出样本方差分别为 $s_1^2=1.40$, $s_2^2=4.38$. 假设电阻值服从正态分布，在显著性水平 $\alpha=0.10$ 下，我们是否可以认为两厂生产的电阻值的方差：

(1) $\sigma_1^2=\sigma_2^2$;

(2) $\sigma_1^2 \leqslant \sigma_2^2$?

第 8 章

方差分析与回归分析简介

方差分析与回归分析都是数理统计中具有广泛应用的重要方法。本章只对它们的最基本部分作简单的介绍。方差分析是根据试验数据,通过分析偏差平方和,合理推断不同因素、同一因素的不同水平对研究对象影响程度的统计分析方法。本章只介绍单因素方差分析。回归分析是研究变量之间非确定的相关关系的统计分析方法。本章重点介绍一元线性回归。

8.1 单因素方差分析

8.1.1 基本概念

在科学试验和生产实践中,影响一事物的因素往往很多.例如在生产过程中,不同的原材料、不同的操作人员以及不同的操作方法等,其中每一个因素的改变都有可能影响产品的数量与质量.有的因素影响大,有的因素影响小;有的因素可以控制,有的因素不能控制.为了使生产过程得以稳定,保证优质、高产,就有必要找出对产品质量有显著影响的那些因素.方差分析就是根据实验结果进行分析,鉴别各因素效应的一种有效的方法.

在前面,我们讨论了两个正态总体均值差的检验问题.我们发现,如果两个总体的方差虽属未知,但若它们相等,就可用 T 检验法来检验其均值是否有显著的差异.那么如果试验中出现了三个或更多个总体,怎样检验它们均值的差异呢? 这就是方差分析所要讨论的问题.

在试验中,我们将要考察的试验结果称为**试验指标**;影响试验指标的条件称为**因素**或**因子**.因素可分为两类,一类是人们可以控制的(可控因素);一类是人们不能控制的.例如,原材料、操作人员、操作方法等是可以控制的,而测量误差、气象条件等一般是难以控制的.以下所说的因素都是指可控因素.

如果在一项试验中只有一个因素在改变,而其他保持不变,这样的试验称为**单因素试验**.处理单因素试验的统计推断问题称为**单因素试验方差分析**.如果多于一个因素在改变,称为**多因素试验**(本章只讨论单因素试验).因素所处的每一个状态或等级称为该因素的**水平**.通常用大写字母 $A,B,C,\cdots$ 表示不同的因素,而用 $A_1,A_2,\cdots,A_r$ 表示因素 A 的 r 个不同的水平.

8.1.2 数学模型

设因素 A 取 r 个不同的水平 $A_1,A_2,\cdots,A_r$，看作有 r 个总体 $X_1,X_2,\cdots,X_r$，它们具有相同的方差. 假定 $X_i\sim N(\mu_i,\sigma^2)$，$i=1,2,\cdots,r$，其中 μ_i,σ^2 均为未知. 在水平 A_i 下，进行 n_i $(n_i\geqslant 2)$ 次独立试验，相当于从总体 X_i 中抽取了容量为 n_i 的样本 $X_{i1},X_{i2},\cdots,X_{in_i}$，$i=1,2,\cdots,r$. 假定这 r 个样本相互独立，于是我们有

$$X_{ij}\sim N(\mu_i,\sigma^2),\quad i=1,2,\cdots,r,j=1,2,\cdots,n_i,$$

且所有的 X_{ij} 相互独立. X_{ij} 就是在水平 A_i 下第 j 次重复试验的试验结果，在实际问题中 X_{ij} 是一个具体数值，而在作统计分析时则将其看成随机变量. 试验结果数据常用表 8-1 表示.

表 8-1

水平 \ 试验结果 \ 重复	1	2	$\cdots$	n_i
A_1	X_{11}	X_{12}	$\cdots$	X_{1n_1}
A_2	X_{21}	X_{22}	$\cdots$	X_{2n_2}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
A_r	X_{r1}	X_{r2}	$\cdots$	X_{rn_r}

令 $\varepsilon_{ij}=X_{ij}-\mu_i$，$j=1,2,\cdots,n_i$，$i=1,2,\cdots,r$，则 ε_{ij} 就是在水平 A_i 下第 j 次重复试验的试验误差，是不可观测的随机变量，称为**随机误差**. 显然有 ε_{ij} 相互独立，且 $\varepsilon_{ij}\sim N(0,\sigma^2)$，$i=1,2,\cdots,r$，$j=1,2,\cdots,n_i$. μ_i 是总体 X_i 的期望值，其意义是水平 A_i 下试验结果数据的理论均值. 此时可把 X_{ij} 表示为

$$X_{ij}=\mu_i+\varepsilon_{ij}. \tag{8.1}$$

于是欲检验的假设为

$$\begin{cases}H_0:\mu_1=\mu_2=\cdots=\mu_r\\ H_1:\mu_1,\mu_2,\cdots,\mu_r\text{ 不全相等}\end{cases}. \tag{8.2}$$

记

$$n=\sum_{i=1}^{r}n_i,\quad \mu=\frac{1}{n}\sum_{i=1}^{r}n_i\mu_i,\quad \alpha_i=\mu_i-\mu,\quad i=1,2,\cdots,r,$$

称 μ 为**理论总均值**；α_i 为水平 A_i 对试验指标的**效应**，简称为 A_i 的效应，它反映了因素 A 的第 i 水平 A_i 对试验指标的作用大小. 显然有 $\sum\limits_{i=1}^{r}n_i\alpha_i=0$. 此时，模型可写为

$$\begin{cases}X_{ij}=\mu+\alpha_i+\varepsilon_{ij}\\ \sum\limits_{i=1}^{r}n_i\alpha_i=0\\ \varepsilon_{ij}\quad(i=1,2,\cdots,r,j=1,2,\cdots,n_i)\text{ 相互独立，且 }\varepsilon_{ij}\sim N(0,\sigma^2)\end{cases}. \tag{8.3}$$

式(8.3)即为单因素试验方差分析的数学模型.

8.1.3 统计分析

1. 总平方和的分解

记

$$\overline{X}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}X_{ij},\quad S_i^2=\frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2,\quad i=1,2,\cdots,r.$$

$\overline{X}_i$ 和 S_i^2 分别称为第 i 个总体的样本均值和样本方差.另记

$$n=\sum_{i=1}^{r}n_i,\quad \overline{X}=\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{n_i}X_{ij}=\frac{1}{n}\sum_{i=1}^{r}n_i\overline{X}_i,$$

$$S^2=\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij}-\overline{X})^2,\quad S_T=nS^2.$$

$\overline{X}$ 称为样本的**总均值**;S^2 称为样本的**总方差**;S_T 称为总的**偏差平方和**,简称**总平方和**.S_T 反映了全体样本 X_{ij} 的波动程度的大小,可将 S_T 作如下分解:

$$\begin{aligned}S_T&=\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij}-\overline{X})^2=\sum_{i=1}^{r}\sum_{j=1}^{n_i}[(X_{ij}-\overline{X}_i)+(\overline{X}_i-\overline{X})]^2\\&=\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2+2\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)(\overline{X}_i-\overline{X})+\sum_{i=1}^{r}\sum_{j=1}^{n_i}(\overline{X}_i-\overline{X})^2.\end{aligned}$$

注意到 $\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)=\sum_{j=1}^{n_i}X_{ij}-n_i\overline{X}_i=0$,所以有

$$\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)(\overline{X}_i-\overline{X})=\sum_{i=1}^{r}(\overline{X}_i-\overline{X})\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)=0.$$

记

$$S_E=\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2=\sum_{i=1}^{r}n_iS_i^2,$$

$$S_A=\sum_{i=1}^{r}\sum_{j=1}^{n_i}(\overline{X}_i-\overline{X})^2=\sum_{i=1}^{r}n_i(\overline{X}_i-\overline{X})^2.$$

于是我们得到总平方和 S_T 的**平方和分解式**

$$S_T=S_E+S_A. \tag{8.4}$$

式中,S_E 称为**误差平方和**(**或组内平方和**).S_E 表示在 A_i 水平下的样本均值与样本值之间的差异,反映了样本内的随机波动,其大小反映了重复试验中随机误差的大小,且 S_E 仅依赖于重复试验中的随机误差.S_A 称为因素 A 的**效应平方和**(**或组间平方和**).S_A 是各水平 A_i 的样本均值 $\overline{X}_i$ 与总均值 $\overline{X}$ 的偏差平方和,反映了样本之间的差异,它是由 A 的不同水平效应的差异以及随机误差引起的.

2. S_E 和 S_A 的统计特性

对于正态总体 $N(\mu_i,\sigma^2)$由式(6.13)有

$$\frac{\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2}{\sigma^2}\sim\chi^2(n_i-1),\quad i=1,2,\cdots,r.$$

注意到 $X_{ij}(i=1,2,\cdots,r,j=1,2,\cdots,n_i)$ 是相互独立的，由 S_E 的定义及 χ^2 分布的可加性可得

$$S_E/\sigma^2 \sim \chi^2(n-r). \tag{8.5}$$

由上式还可得到

$$E(S_E/(n-r)) = \sigma^2. \tag{8.6}$$

另外可以证明：

(1) $E(S_A) = (r-1)\sigma^2 + \sum_{i=1}^{r} n_i \alpha_i^2$；

(2) S_A 与 S_E 相互独立，且当 H_0 成立时有

$$S_A/\sigma^2 \sim \chi^2(r-1), \tag{8.7}$$

$$E(S_A) = (r-1)\sigma^2. \tag{8.8}$$

3. 假设(8.2)的检验法

由式(8.5)和式(8.7)以及 F 分布的定义可知，当 H_0 成立时有

$$\frac{S_A/\sigma^2}{r-1}\Bigg/\frac{S_E/\sigma^2}{n-r} \sim F(r-1,n-r),$$

于是可以建立检验假设(8.2)的检验法. 取

$$F = \frac{S_A/(r-1)}{S_E/(n-r)} \tag{8.9}$$

为检验统计量. 由前述的分析可知，对于给定的显著性水平 α，假设(8.2)的拒绝域为

$$F = \frac{S_A/(r-1)}{S_E/(n-r)} \geqslant F_\alpha(r-1,n-r). \tag{8.10}$$

即当 $F \geqslant F_\alpha(r-1,n-r)$ 时，拒绝 H_0，此时，不同水平的差异显著；否则，接受 H_0.

上述的检验法称为**方差分析法**，其分析结果常列成表 8-2 的形式，称为**方差分析表**.

表 8-2

方差来源	平方和	自由度	均方	F 值
因素 A	S_A	$r-1$	$\overline{S}_A = S_A/(r-1)$	$F=\dfrac{\overline{S}_A}{\overline{S}_E}$
误差	S_E	$n-r$	$\overline{S}_E = S_E/(n-r)$	
总和	S_T	$n-1$		

注：通常为简化计算过程，提高计算精度，可按下列顺序计算 S_T, S_A, S_E.

$$S_T = \sum_{i=1}^{r}\sum_{j=1}^{n_i} X_{ij}^2 - \frac{1}{n}\left(\sum_{i=1}^{r}\sum_{j=1}^{n_i} X_{ij}\right)^2,$$

$$S_A = \sum_{i=1}^{r}\frac{1}{n_i}\left(\sum_{j=1}^{n_i} X_{ij}\right)^2 - \frac{1}{n}\left(\sum_{i=1}^{r}\sum_{j=1}^{n_i} X_{ij}\right)^2,$$

$$S_E = S_T - S_A.$$

例 1 某灯泡厂用 4 种不同的配料方案制成的灯丝生产了四批灯泡，在每批灯泡中随机抽取若干个灯泡，测量其使用寿命(单位：h)，数据见表 8-3.

表 8-3 h

灯丝	灯泡							
	1	2	3	4	5	6	7	8
甲	1600	1610	1650	1680	1700	1700	1780	
乙	1500	1640	1400	1700	1750			
丙	1640	1550	1600	1620	1640	1600	1740	1800
丁	1510	1520	1530	1570	1640	1680		

试问这 4 种灯丝生产的灯泡的使用寿命有无显著差异.($\alpha=0.1$)

解 本题中灯泡的使用寿命为试验指标,灯丝的配料为因素,4 种不同的配料方案为 4 个水平,故本题是一个单因素四水平的试验.于是有

原假设 $H_0:\mu_1=\mu_2=\mu_3=\mu_4$;

备择假设 $H_1:\mu_1,\mu_2,\mu_3,\mu_4$ 不全相等.

由表 8-3 得 $r=4,n_1=7,n_2=5,n_3=8,n_4=6,n=26$,又 $F_{0.1}(3,22)=2.35$.代入相关公式经计算可得方差分析表(见表 8-4).

因为 $1.638<2.35$,故接受 H_0,即认为用 4 种不同的灯丝配料生产的灯泡其平均使用寿命无显著性差异.从而,灯泡厂在选择灯丝配料方案时就可以从降低成本等方面来考虑,而无须过多考虑使用寿命方面的问题.

表 8-4

方差来源	平方和	自由度	均方	F 值
灯丝配料	39 776.4	3	13 258.8	1.638
误差	178 089	22	8095	
总和	217 865.4	25		

例 2 今有 6 种不同的农药,在相同的条件下分别进行杀虫试验,试验结果见表 8-5.问:杀虫率是否因农药的不同而有显著的差异?($\alpha=0.01$)

表 8-5

杀虫率/% 农药 / 试验号	A_1	A_2	A_3	A_4	A_5	A_6
1	87	90	56	55	92	75
2	85	88	62	48	99	72
3	80	87			95	81
4		94			91	

解 本题是一个单因素六水平的试验.于是有

原假设 $H_0:\mu_1=\mu_2=\cdots=\mu_6$;

备择假设 $H_1:\mu_1,\mu_2,\cdots,\mu_6$ 不全相等.

由表 8-5 得 $r=6,n_1=3,n_2=4,n_3=2,n_4=2,n_5=4,n_6=3,n=18$.代入相关公式经计

算可得方差分析表(见表 8-6):

表 8-6

方差来源	平方和	自由度	均方	F 值
农药	3794.5	5	758.9	51.17
误差	178	12	14.83	
总和	3972.5	17		

因为 $51.17>F_{0.01}(5,12)=5.06$,故拒绝 H_0,即认为不同农药的杀虫率有显著差异.由表 8-5 可知,第五种农药的杀虫率最高,其平均杀虫率高达 94.25%.

另外,在实际问题中,影响试验结果的因素往往不止一个,这时要考虑每个不同因素的作用,就要用到多因素试验的方差分析,相关内容读者可参考其他有关书籍.

8.2 一元线性回归

8.2.1 回归的含义

在客观世界中普遍存在的变量之间的关系一般来说可分为确定性的与非确定性的两种.确定性关系是指变量之间的关系可以用函数关系来表达,如圆的周长 l 与半径 r 之间的关系为 $l=2\pi r$.非确定性的关系也可称为**相关关系**,这种关系无法用一个精确的数学式子来表示.例如人的身高与体重之间的关系、学习成绩与学习时间之间的关系、收入与智商之间的关系等,都是相关关系.这些相关关系中的变量可以是随机的,也可以是非随机的.我们要用到的变量主要有两种类型.一种变量相当于函数关系中的自变量,该变量能够赋予一个需要的值(如试验的温度),或能取到一个可以观测但不能人为控制的值(如人的身高).这种变量称为**预报变量**(或**自变量**).预报变量的变化能影响另一些变量的变化,这样的变量称为**响应变量**(或**因变量**).预报变量与响应变量间无明显的界线,往往与所考虑的实际问题有关.由预报变量(非随机变量)来估计或预测某一响应变量(随机变量)时,所进行的统计分析称为**回归分析**.回归分析就是研究相关关系的一种数学工具.

回归分析用来研究预报变量的变化对响应变量的变化的影响程度,其目的在于根据已知的预报变量的变化来估计或预测响应变量的变化情况.回归(regression)一词是由英国生物统计学家 Francis Galton 于 1885 年提出的.他发现:如果父辈身材高大,则后代的身材要比父辈的身材矮一些;如果父辈身材矮小,则后代的身材要比父辈的身材高大一些.即后代的身高有向平均值靠拢的趋向,在两端的高度的后代,其高度有向中心回归的趋势,离中心越远,所受回归的压力越大.因此,他用回归一词来描述后代的身高与父辈的身高的这种关系.尔后,英国著名统计学家 K. Pearson 等人搜集了上千个家庭成员的身高数据,分析出儿子的身高 y 与父亲的身高 x 间的关系,大致可以表示为 $y=0.516x+33.73$(单位:in),从而进一步证实了 Galton 的"回归定律".

8.2.2 一元线性回归

只有一个预报变量的回归分析称为**一元回归分析**.多于一个预报变量的回归分析称为

多元回归分析. 当变量间具有线性关系时,相应的回归分析称为**线性回归分析**.

对于预报变量 x 的每一个值,响应变量 Y 的取值是随机的,但它的总体均值 $E(Y)$是唯一的,它是 x 的函数,记为 $f(x)$(即 $E(Y)=f(x)$). 又 $Y=E(Y)+\varepsilon$,其中 ε 是随机变量,$E(\varepsilon)=0$,称 $E(Y)$为 Y 关于 x 的**回归函数**. 方程$\hat{y}=f(x)$称为**回归方程**. 相应的图形称为**回归曲线**. 特别地,当回归曲线为直线时,称为**回归直线**.

函数 $f(x)$的具体形式可通过样本进行估计. 对于 x 的一组值 $x_1,x_2,\cdots,x_n$,作独立试验,得到关于 Y 的 n 个观察结果 $y_1,y_2,\cdots,y_n$. 于是有 n 对观察值:

$$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n).$$

这 n 对结果就是容量为 n 的样本,我们要解决的问题是如何利用样本估计 $f(x)$. 通常可以根据上述的观察值,在直角坐标系中描出相应的点,所得到的图形叫**散点图**. 利用散点图可以粗略地看出 Y 与 x 的关系,从而推测出 $f(x)$的形式,并从这种形式出发,作进一步的分析.

1. 一元线性回归模型

这里我们只考虑回归函数 $f(x)$是 x 的线性函数的情况. 现假设对于 x 在某个区间内的每一个值都有可观测的随机变量 $Y=a+bx+\varepsilon$,其中 a,b 为未知参数,ε 是不可观测的随机变量,且满足 $E(\varepsilon)=0,D(\varepsilon)=\sigma^2$($\sigma^2$ 未知,且 $\sigma^2<+\infty$). 于是可得到下面的模型:

$$\begin{cases} Y=a+bx+\varepsilon \\ E(\varepsilon)=0, \quad D(\varepsilon)=\sigma^2 \end{cases}. \tag{8.11}$$

式(8.11)称为**一元线性回归模型**.

如果能够求出式(8.11)中 a,b 的估计$\hat{a},\hat{b}$,取$\hat{y}=\hat{a}+\hat{b}x$ 作为 $f(x)=a+bx$ 的估计,此时,称方程$\hat{y}=\hat{a}+\hat{b}x$ 为 Y 关于 x 的**一元线性回归方程**. 未知参数 a,b 称为**回归系数**,预报变量 x 也称为**回归变量**. $\hat{y}$表示$\hat{a},\hat{b}$确定之后,对于每个给定的 x,与之相应的 Y 的**预报值**(**或拟合值**、**回归值**). 方程$\hat{y}=\hat{a}+\hat{b}x$ 称为**预报方程**(**或拟合方程**、**回归方程**),对应的直线称为**回归直线**(**或拟合直线**).

2. 最小二乘估计

现利用最小二乘法来估计模型式(8.11)中的未知参数 a,b. 假设有 n 组独立观察值$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$,由式(8.11)可得

$$\begin{cases} y_i=a+bx_i+\varepsilon_i \\ E(\varepsilon_i)=0, \quad D(\varepsilon_i)=\sigma^2, \quad \varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n \text{ 相互独立} \end{cases}. \tag{8.12}$$

令

$$Q(a,b)=\sum_{i=1}^{n}(y_i-a-bx_i)^2, \tag{8.13}$$

$Q(a,b)$称为偏离真实直线的**偏差平方和**.

用最小二乘法估计 a,b,就是使 $Q(\hat{a},\hat{b})=\min Q(a,b)$.

分别求 $Q(a,b)$关于 a,b 的偏导数,并令其等于 0,可得

$$\begin{cases}\dfrac{\partial Q}{\partial a}=-2\sum_{i=1}^{n}(y_i-a-bx_i)=0\\ \dfrac{\partial Q}{\partial b}=-2\sum_{i=1}^{n}(y_i-a-bx_i)x_i=0\end{cases},$$

得到方程组

$$\begin{cases}na+\left(\sum_{i=1}^{n}x_i\right)b=\sum_{i=1}^{n}y_i\\ \left(\sum_{i=1}^{n}x_i\right)a+\left(\sum_{i=1}^{n}x_i^2\right)b=\sum_{i=1}^{n}x_iy_i\end{cases}.$$

上式称为**正规方程组**.解该方程组可得

$$\begin{cases}\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}\\ \hat{a}=\bar{y}-\hat{b}\bar{x}\end{cases},\tag{8.14}$$

其中,$\bar{x}=\dfrac{1}{n}\sum_{i=1}^{n}x_i,\bar{y}=\dfrac{1}{n}\sum_{i=1}^{n}y_i$.

此时,所求的一元线性回归方程为

$$\hat{y}=\hat{a}+\hat{b}x.\tag{8.15}$$

若将$\hat{a}=\bar{y}-\hat{b}\bar{x}$ 代入式(8.15),可得

$$\hat{y}-\bar{y}=\hat{b}(x-\bar{x}).\tag{8.16}$$

式(8.16)表明,回归直线通过散点图的几何中心$(\bar{x},\bar{y})$.

为了便于记忆和书写,常引入下述记号:

$$\begin{cases}S_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-\dfrac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2\\ S_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-\dfrac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2=\sum_{i=1}^{n}y_i^2-n\bar{y}^2\\ S_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}x_iy_i-\dfrac{1}{n}\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)=\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}\end{cases},$$

此时,可将式(8.14)改写成

$$\begin{cases}\hat{b}=\dfrac{S_{xy}}{S_{xx}}\\ \hat{a}=\bar{y}-\hat{b}\bar{x}\end{cases}.\tag{8.17}$$

例 1 为研究温度对某化学过程的产量的影响,作试验得到数据如表 8-7 所示.

表 8-7

温度 x/℃	−5	−4	−3	−2	−1	0	1	2	3	4	5
产量 y	1	5	4	7	10	8	9	13	14	13	18

求：Y 关于 x 的回归方程.

解 利用表 8-7 中的数据及式(8.17)，经计算可得

$$n=11,\quad \bar{x}=0,\quad \sum_{i=1}^{11} x_i^2=110,\quad \bar{y}=9.273,\quad \sum_{i=1}^{11} x_i y_i=158,$$

$$S_{xx}=110,\quad S_{xy}=158,\quad \hat{b}=\frac{S_{xy}}{S_{xx}}=1.436,\quad \hat{a}=\bar{y}-\hat{b}\bar{x}=9.273.$$

所以回归方程为

$$\hat{y}=9.273+1.436x.$$

3. σ^2的点估计

对于每个 x_i，由回归方程 $\hat{y}=\hat{a}+\hat{b}x$ 有 $\hat{y}_i=\hat{a}+\hat{b}x_i$，称 $y_i-\hat{y}_i$ 为 x_i 处的**残差**，并称平方和

$$Q_e=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\sum_{i=1}^{n}(y_i-\hat{a}-\hat{b}x_i)^2 \tag{8.18}$$

为**残差平方和**. 式(8.18)经化简有 $Q_e=S_{yy}-\hat{b}^2S_{xx}$. 另外可证明

$$\frac{Q_e}{\sigma^2}\sim\chi^2(n-2), \tag{8.19}$$

于是有 $E\left(\frac{Q_e}{\sigma^2}\right)=n-2$，即 $E\left(\frac{Q_e}{n-2}\right)=\sigma^2$，从而可得

$$\hat{\sigma}^2=\frac{Q_e}{n-2}. \tag{8.20}$$

它是 σ^2 的无偏估计. 式(8.20)即为 σ^2 的点估计.

4. 线性假设的显著性检验

我们所求的一元线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$，是在线性假设 $Y=a+bx+\varepsilon$ 的前提下得出的. 该线性回归方程是否符合实际，需通过假设检验的方法来进行判断. 注意到若 b 等于 0，则 Y 与 x 之间不存在线性关系，故问题的实质是回归系数 b 是否为 0. 因此，可检验如下的假设：

$$\begin{cases} H_0: b=0 \\ H_1: b\neq 0 \end{cases}. \tag{8.21}$$

若在模型式(8.11)中，进一步假定 $\varepsilon\sim N(0,\sigma^2)$，可以证明 $\hat{b}\sim N(b,\sigma^2/S_{xx})$. 对 $\hat{b}$ 进行标准化可得

$$\frac{\hat{b}-b}{\sqrt{\sigma^2/S_{xx}}}\sim N(0,1), \tag{8.22}$$

且 $\hat{b}$ 与 Q_e 相互独立. 由式(8.19)、式(8.20)、式(8.22)及 t 分布的定义，有

$$\frac{\hat{b}-b}{\sqrt{\sigma^2/S_{xx}}}\Bigg/\sqrt{\frac{(n-2)\hat{\sigma}^2}{\sigma^2}\Big/(n-2)}\sim t(n-2),$$

即

$$\frac{\hat{b}-b}{\hat{\sigma}}\sqrt{S_{xx}}\sim t(n-2). \tag{8.23}$$

于是取检验统计量为

$$T=\frac{\hat{b}}{\hat{\sigma}}\sqrt{S_{xx}}. \tag{8.24}$$

对于给定的显著水平 α，H_0 的拒绝域为

$$|T|=\left|\frac{\hat{b}}{\hat{\sigma}}\right|\sqrt{S_{xx}}\geqslant t_{\frac{\alpha}{2}}(n-2). \tag{8.25}$$

若拒绝 H_0，说明回归效果显著；而接受 H_0，说明回归效果不显著.

当回归效果显著时，还可进一步计算回归系数 b 的置信度为 $1-\alpha$ 的置信区间. 具体结果如下：

$$(\underline{b},\bar{b})=\left(\hat{b}-t_{\frac{\alpha}{2}}(n-2)\frac{\hat{\sigma}}{\sqrt{S_{xx}}},\hat{b}+t_{\frac{\alpha}{2}}(n-2)\frac{\hat{\sigma}}{\sqrt{S_{xx}}}\right). \tag{8.26}$$

例 2 检验例 1 的回归方程的效果是否显著($\alpha=0.01$). 若显著，则求出 b 的置信度为 0.95 的置信区间.

解 由式(8.24)得

$$t=\frac{\hat{b}}{\hat{\sigma}}\sqrt{S_{xx}}=\frac{1.436}{Q_e/(n-2)}\sqrt{110}=9.7798,$$

而 $t_{\frac{\alpha}{2}}(n-2)=t_{0.005}(9)=3.2498$. 由于 $9.7798>3.2498$，拒绝 H_0，因而，可认为线性回归效果显著.

由式(8.26)可得 b 的置信度为 0.95 的置信区间为

$$(\underline{b},\bar{b})=\left(0.483-2.306\times\sqrt{\frac{0.934}{8250}},0.483+2.306\times\sqrt{\frac{0.934}{8250}}\right)$$
$$=(0.458,0.508).$$

5. 利用回归方程进行预测与控制

当回归效果显著时，可以利用回归方程进行预测与控制. 所谓预测问题就是对给定的 x 值预测相应的 y 的值；而所谓控制问题是指通过控制 x 的值以便把 y 的值控制在指定的范围之内.

1) 预测

设 x 与 y 满足线性回归模型(8.11)，且 $\varepsilon\sim N(0,\sigma^2)$. 若 x_0 表示 x 的某个给定的值，由回归方程的意义，可用回归值 $\hat{y}_0=\hat{a}+\hat{b}x_0$ 作为 y_0 的预测值，且 $\hat{y}_0$ 是 Ey_0 的无偏估计.

可证明

$$\hat{y}_0\sim N\left(a+bx_0,\left[\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]\sigma^2\right),$$

故

$$y_0-\hat{y}_0\sim N\left(0,\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]\sigma^2\right).$$

经标准化可得

$$U=\frac{y_0-\hat{y}_0}{\sigma\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}}}\sim N(0,1).$$

由式(8.19)和式(8.20)可得

$$V=\frac{(n-2)\hat{\sigma}^2}{\sigma^2}\sim\chi^2(n-2).$$

由 $y_0,\hat{y}_0,\hat{\sigma}^2$ 是相互独立的,以及 t 分布的定义,可知

$$T=\frac{y_0-\hat{y}_0}{\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}}}\sim t(n-2).$$

对于置信度 $1-\alpha$,有

$$P\{|T|<t_{\frac{\alpha}{2}}(n-2)\}=1-\alpha,$$

即

$$P\{\hat{y}_0-\delta(x_0)<y_0<\hat{y}_0+\delta(x_0)\}=1-\alpha,$$

其中

$$\delta(x_0)=\hat{\sigma}t_{\frac{\alpha}{2}}(n-2)\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}}. \tag{8.27}$$

于是得到 y_0 的置信度为 $1-\alpha$ 的预测区间为

$$(\hat{y}_0-\delta(x_0),\hat{y}_0+\delta(x_0)). \tag{8.28}$$

由 $\delta(x_0)$的定义可知,残差平方和 Q_e 越小,预测区间的长度就越小,即预测越精确.另外,对于样本值和置信度给定的情形,x_0 越接近 $\bar{x}$,预测精度也越高.

例 3 求例 1 中当 $x_0=3$ 时,y_0 的预测值和置信度为 0.95 的预测区间.

解 由例 1 可得回归方程为

$$\hat{y}=9.273+1.436x,$$

所以,当 $x_0=3$ 时 y_0 的预测值为

$$\hat{y}_0=9.273+1.436\times3=13.581.$$

当 $x_0=3$ 时,由式(8.27)可得

$$\delta(x_0)=\hat{\sigma}t_{\frac{\alpha}{2}}(n-2)\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}}=3.7631,$$

所以由式(8.28)得 y_0 的置信度为 0.95 的预测区间为(9.818,17.344).

2) 控制

控制问题是预测的反问题,相当于要使 y 以置信度 $1-\alpha$ 在指定区间(y_1,y_2)内取值,x 应控制在什么范围内.即求出区间(x_1,x_2),使当 $x_1<x<x_2$ 时,对应的 y 的值以概率 $1-\alpha$ 落在指定的区间(y_1,y_2)内.

本书仅讨论样本容量较大,且 x 在 $\bar{x}$ 附近的情形.此时利用

$$t_{\frac{\alpha}{2}}(n-2)\approx z_{\frac{\alpha}{2}},\quad 1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\approx1$$

来简化求解过程.经计算可得 x 的控制区间为

$$\left((y_1-\hat{a}+z_{\frac{\alpha}{2}}\hat{\sigma})/\hat{b},\quad(y_2-\hat{a}-z_{\frac{\alpha}{2}}\hat{\sigma})/\hat{b}\right). \tag{8.29}$$

注:欲实现上述控制,应使区间(y_1,y_2)的长度大于 $2z_{\frac{\alpha}{2}}\hat{\sigma}$.

8.2.3 常用非线性回归的线性化方法

实际问题中,预测变量与响应变量间不一定呈线性的相关关系.在某些特殊情况下,可通过对预测变量和响应变量作适当的变换,将非线性的相关关系转化为线性相关关系,以便用线性回归的方法进行处理.下面以四种最常见的情形为例加以说明.

1. 双曲线

对于双曲线

$$\frac{1}{y}=a+\frac{b}{x},$$

作变换 $u=\frac{1}{y}$,$v=\frac{1}{x}$,可得线性函数

$$u=a+bv.$$

2. 幂函数

对于幂函数

$$y=ax^b,$$

作变换 $u=\ln y$,$v=\ln x$,$c=\ln a$,可得线性函数

$$u=c+bv.$$

3. 指数函数

对于指数函数

$$y=a\mathrm{e}^{bx},$$

作变换 $u=\ln y$,$c=\ln a$,可得线性函数

$$u=c+bx.$$

4. 对数函数

对于对数函数

$$y=a+b\ln x,$$

作变换 $v=\ln x$,可得线性函数

$$y=a+bv.$$

在实际应用中,一般是根据给出的 n 组数据 (x_1,y_1),(x_2,y_2),…,(x_n,y_n),在平面上标出 n 个点,根据这 n 个点所呈现的形状,与直线或以上几种常见的曲线进行比较,选择直线或曲线进行拟合.

例 4 人的主动脉压与主动脉容积的关系如表 8-8 所示.

表 8-8

容积 V/ml	10	25	50	75	100	125	150
脉压 P/mmHg	2	6	18	38	62	93	138

试用指数函数来拟合主动脉压与主动脉容积的关系曲线,求出相应的回归方程.

解 设指数函数为 $P=a\mathrm{e}^{bV}$,并作变换 $u=\ln P$,$c=\ln a$,可得线性函数

$$u=c+bV.$$

由表 8-8 的数据可得新的数据如下：

V	10	25	50	75	100	125	150
$u=\ln P$	0.6931	1.7981	2.8904	3.6376	4.1271	4.5850	4.9273

将上述新的数据代入式(8.17)，经计算可得

$$c=1.0376,\quad b=0.0288,$$

所以得到相应的回归方程为

$$P=2.822e^{0.0288V}.$$

在实际问题中，影响结果 Y 的因素往往不止一个. 此时，预报变量大于一个，该问题属于多元回归问题. 感兴趣的读者可参考其他有关书籍.

小　　结

本章简要介绍了方差分析与回归分析中的基本内容。通过对单因素方差分析和一元线性回归的学习，使学生了解如何运用统计分析方法解决实际应用问题，并初步培养学生将实际问题归结为数学问题的建模能力。

基本要求：

(1) 了解单因素试验的方差分析的数学模型，了解平方和的分解.

(2) 了解 S_E 和 S_A 的统计特性，能够运用方差分析表处理简单的单因素方差分析问题.

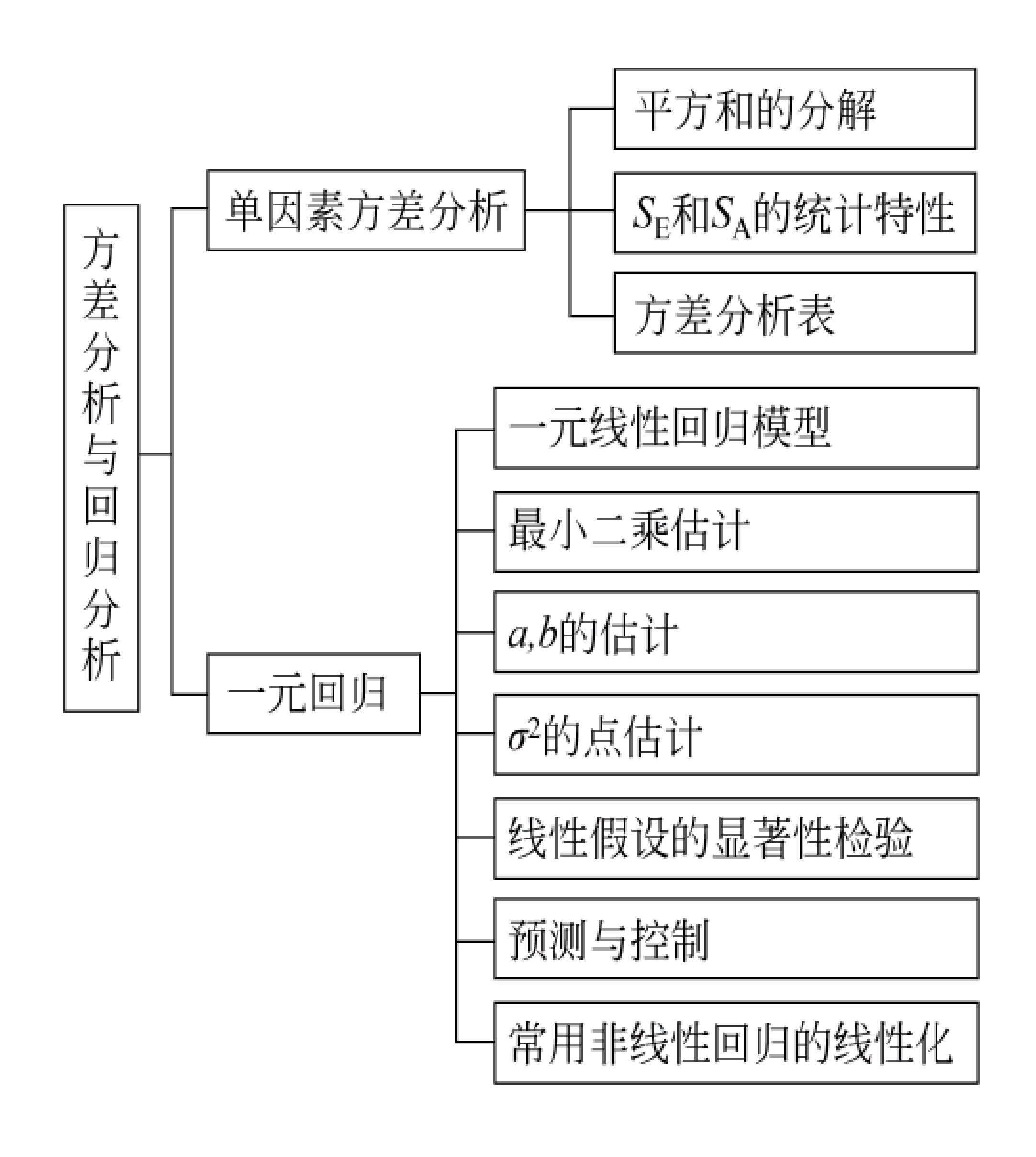
(3) 了解回归的概念，了解一元线性回归模型，会用最小二乘法计算 $\hat{a},\hat{b}$ 的估计.

(4) 了解线性假设的显著性检验，会利用回归方程进行简单的预测与控制，了解常用非线性回归的线性化方法.

重点：平方和的分解，方差分析表，一元线性回归方程的求解，最小二乘法.

难点：运用方差分析表处理单因素方差分析问题，一元线性回归方程的求解及相关计算.

本章知识结构：



习 题 8

1. 设有 3 台机器生产规格相同的薄钢板.现从生产出的钢板中各取 5 块,测出其厚度值,如表 8-9 所示.

表 8-9 mm

机器	厚度测量值				
Ⅰ	2.36	2.38	2.48	2.45	2.43
Ⅱ	2.57	2.53	2.55	2.54	2.61
Ⅲ	2.58	2.64	2.59	2.67	2.62

设各测量值服从同方差的正态分布,试分析各机器生产的薄钢板的厚度有无显著差异.($\alpha=0.05$)

2. 考察 4 种不同类型的电路对计算器的响应时间的影响,测得数据如表 8-10 所示.设各测量值服从同方差的正态分布,试分析各类型的电路对计算器的响应时间有无显著差异.($\alpha=0.05$)

表 8-10 ms

电路类型	响应时间				
Ⅰ	19	22	20	18	15
Ⅱ	20	21	33	27	40
Ⅲ	16	15	18	26	17
Ⅳ	18	22	19		

3. 某学院有 3 个班,进行了一次高等数学考试.从各班随机抽取部分学生,记录其数学成绩如表 8-11 所示.设各测量值服从同方差的正态分布,试分析各班成绩有无显著差异.($\alpha=0.05$)

表 8-11 分

班级	成绩
1	73 66 89 60 82 45 93 80 36 77 43 73
2	88 77 78 31 48 78 91 62 51 76 85 96 74 80 56
3	87 68 41 79 59 71 56 68 91 53 79 71 15

4. 测量 4 种不同类型外壳的彩色显像管的传导率,得到传导率的观察值如表 8-12 所示.

表 8-12

类型 1	143	141	150	146
类型 2	152	144	137	143
类型 3	134	136	133	129
类型 4	129	128	134	129

试分析外壳类型对传导率有无显著影响.($\alpha=0.05$)

5. 某河的径流量 y 与该地区的降雨量 x 有关,测量数据如表 8-13 所示.设径流量 y 是

正态变量，方差与 x 无关.

表 8-13

x_i	110	184	145	122	165	143	78	129	62	130	168
y_i	25	81	36	33	70	54	20	44	1.4	41	75

求：(1) 径流量 y 关于降雨量 x 的一元线性回归方程；

(2) σ^2 的估计值；

(3) $x_0=155$ 时，y_0 的预测值.

6. 为研究某一化学反应过程中温度 x(单位：℃)对产品得率 y(%)的影响，测得数据如表 8-14 所示.求：(1) y 关于 x 的一元线性回归方程；(2) σ^2 的无偏估计；(3)检验回归效果是否显著.($\alpha=0.05$)

表 8-14

x_i	100	110	120	130	140	150	160	170	180	190
y_i	45	51	54	61	66	70	74	78	85	89

7. 炼铝厂测得铝的硬度 x 与抗张强度 y 的数据如表 8-15 所示.

表 8-15

x_i	68	53	70	84	60	72	51	83	70	64
y_i	288	298	349	343	290	354	283	324	340	286

(1) 求抗张强度 y 对硬度 x 的回归方程；

(2) 检验回归方程的显著性($\alpha=0.05$)；

(3) 求 y 在 $x=65$ 处的预测区间(置信度为 0.95).

8. 某建材实验室在作陶粒混凝土强度试验中，考察每立方米混凝土的水泥用量 x(单位：kg)对 28 天后的混凝土抗压强度 y(单位：kg/cm^2)的影响.测得如表 8-16 所示数据.

表 8-16

x_i	150	160	170	180	190	200	220	230	240	250	260
y_i	56.9	58.3	61.6	64.6	71.3	74.1	77.4	80.2	82.6	86.4	89.7

(1) 求 y 对 x 的线性回归方程，并求每立方米混凝土中增加 1 kg 水泥时，可提高的抗压强度是多少；

(2) 检验线性回归效果的显著性($\alpha=0.05$)；

(3) 求 $x_0=22.5$ kg 时，y_0 的预测值及预测区间(置信度为 0.95).

补充与提高

9. 由于钢水对耐火材料的浸蚀，出钢时所用的盛钢水的钢包的容积不断增大.我们希望找出使用次数 x 与增大的容积 Y 之间的关系.试验数据列于表 8-17. (1)根据试验

数据做出散点图；(2)用双曲线拟合数据，求出相应的回归方程；(3)检验回归效果是否显著($\alpha=0.01$).

表　8-17

x	2	3	4	5	6	7	8	9	10
y	6.42	8.20	9.58	9.50	9.70	10.00	9.83	9.99	10.49
x	11	12	13	14	15	16			
y	10.59	10.60	10.80	10.60	10.90	10.76			

附录 A 常用概率统计表

附表 A.1 泊松分布表

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}$$

k \ λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.5	2.0	2.5	3.0
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679	0.2231	0.1353	0.0821	0.0498
1	0.0905	0.1637	0.2223	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659	0.3679	0.3347	0.2707	0.2052	0.1494
2	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1216	0.1438	0.1647	0.1839	0.2510	0.2707	0.2565	0.2240
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494	0.0613	0.1255	0.1805	0.2138	0.2240
4		0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111	0.0153	0.0471	0.0902	0.1336	0.1681
5				0.0001	0.0002	0.0003	0.0007	0.0012	0.0020	0.0031	0.0141	0.0361	0.0668	0.1008
6							0.0001	0.0002	0.0003	0.0005	0.0035	0.0120	0.0278	0.0504
7										0.0001	0.0008	0.0034	0.0099	0.0216
8											0.0002	0.0009	0.0031	0.0081
9												0.0002	0.0009	0.0027
10													0.0002	0.0008
11													0.0001	0.0002
12														0.0001

续表

k \ λ	3.5	4.0	4.5	5	6	7	8	9	10	11	12	13	14	15
0	0.0302	0.0183	0.0111	0.0067	0.0025	0.0009	0.0003	0.0001						
1	0.1057	0.0733	0.0500	0.0337	0.0149	0.0064	0.0027	0.0011	0.0004	0.0002	0.0001			
2	0.1850	0.1465	0.1125	0.0842	0.0446	0.0223	0.0107	0.0050	0.0023	0.0010	0.0004	0.0002	0.0001	
3	0.2158	0.1954	0.1687	0.1404	0.0892	0.0521	0.0286	0.0150	0.0076	0.0037	0.0018	0.0008	0.0004	0.0002
4	0.1888	0.1954	0.1898	0.1755	0.1339	0.0912	0.0573	0.0337	0.0189	0.0102	0.0053	0.0027	0.0013	0.0006
5	0.1322	0.1563	0.1708	0.1755	0.1606	0.1277	0.0916	0.0607	0.0378	0.0224	0.0127	0.0071	0.0037	0.0019
6	0.0771	0.1042	0.1281	0.1462	0.1606	0.1490	0.1221	0.0911	0.0631	0.0411	0.0255	0.0151	0.0087	0.0048
7	0.0385	0.0595	0.0824	0.1044	0.1377	0.1490	0.1396	0.1171	0.0901	0.0646	0.0437	0.0281	0.0174	0.0104
8	0.0169	0.0298	0.0463	0.0653	0.1033	0.1304	0.1396	0.1318	0.1126	0.0888	0.0655	0.0457	0.0304	0.0195
9	0.0065	0.0132	0.0232	0.0363	0.0688	0.1014	0.1241	0.1318	0.1251	0.1085	0.0874	0.0660	0.0473	0.0324
10	0.0023	0.0053	0.0104	0.0181	0.0413	0.0710	0.0993	0.1186	0.1251	0.1194	0.1048	0.0859	0.0663	0.0486
11	0.0007	0.0019	0.0043	0.0082	0.0225	0.0452	0.0722	0.0970	0.1137	0.1194	0.1144	0.1015	0.0843	0.0663
12	0.0002	0.0006	0.0015	0.0034	0.0113	0.0264	0.0481	0.0728	0.0948	0.1094	0.1144	0.1099	0.0984	0.0828
13	0.0001	0.0002	0.0006	0.0013	0.0052	0.0142	0.0296	0.0504	0.0729	0.0926	0.1056	0.1099	0.1061	0.0956
14		0.0001	0.0002	0.0005	0.0023	0.0071	0.0169	0.0324	0.0521	0.0728	0.0905	0.1021	0.1061	0.1025
15			0.0001	0.0002	0.0009	0.0033	0.0090	0.0194	0.0347	0.0533	0.0724	0.0885	0.0989	0.1025
16				0.0001	0.0003	0.0015	0.0045	0.0109	0.0217	0.0367	0.0543	0.0719	0.0865	0.0960
17					0.0001	0.0006	0.0021	0.0058	0.0128	0.0237	0.0383	0.0551	0.0713	0.0847
18						0.0002	0.0010	0.0029	0.0071	0.0145	0.0255	0.0397	0.0554	0.0706
19						0.0001	0.0004	0.0014	0.0037	0.0084	0.0161	0.0272	0.0408	0.0557
20							0.0002	0.0006	0.0019	0.0046	0.0097	0.0177	0.0286	0.0418
21							0.0001	0.0003	0.0009	0.0024	0.0055	0.0109	0.0191	0.0299
22								0.0001	0.0004	0.0013	0.0030	0.0065	0.0122	0.0204
23									0.0002	0.0006	0.0016	0.0036	0.0074	0.0133
24									0.0001	0.0003	0.0008	0.0020	0.0043	0.0083
25										0.0001	0.0004	0.0011	0.0024	0.0050
26											0.0002	0.0005	0.0013	0.0029
27											0.0001	0.0002	0.0007	0.0017
28												0.0001	0.0003	0.0009
29													0.0002	0.0004
30													0.0001	0.0002
31														0.0001

续表

λ=20						λ=30					
k	p	k	p	k	p	k	p	k	p	k	p
5	0.0001	20	0.0889	35	0.0007	10		25	0.0511	40	0.0139
6	0.0002	21	0.0846	36	0.0004	11		26	0.0590	41	0.0102
7	0.0006	22	0.0769	37	0.0002	12	0.0001	27	0.0655	42	0.0073
8	0.0013	23	0.0669	38	0.0001	13	0.0002	28	0.0702	43	0.0051
9	0.0029	24	0.0557	39	0.0001	14	0.0005	29	0.0727	44	0.0035
10	0.0058	25	0.0446			15	0.0010	30	0.0727	45	0.0023
11	0.0106	26	0.0343			16	0.0019	31	0.0703	46	0.0015
12	0.0176	27	0.0254			17	0.0034	32	0.0659	47	0.0010
13	0.0271	28	0.0183			18	0.0057	33	0.0599	48	0.0006
14	0.0382	29	0.0125			19	0.0089	34	0.0529	49	0.0004
15	0.0517	30	0.0083			20	0.0134	35	0.0453	50	0.0002
16	0.0646	31	0.0054			21	0.0192	36	0.0378	51	0.0001
17	0.0760	32	0.0034			22	0.0261	37	0.0306	52	0.0001
18	0.0844	33	0.0021			23	0.0341	38	0.0242		
19	0.0889	34	0.0012			24	0.0426	39	0.0186		
λ=40						λ=50					
k	p	k	p	k	p	k	p	k	p	k	p
18	0.0001	35	0.0485	55	0.0043	26	0.0001	43	0.0363	60	0.0201
19	0.0001	36	0.0539	56	0.0031	27	0.0001	44	0.0412	61	0.0165
20	0.0002	37	0.0583	57	0.0022	28	0.0002	45	0.0458	65	0.0063
21	0.0004	38	0.0614	58	0.0015	29	0.0004	46	0.0498	66	0.0048
22	0.0007	39	0.0629	59	0.0010	30	0.0007	47	0.0530	67	0.0036
23	0.0012	40	0.0629	60	0.0007	31	0.0011	48	0.0552	68	0.0026
24	0.0019	41	0.0614	61	0.0005	32	0.0017	49	0.0564	69	0.0019
25	0.0031	42	0.0585	62	0.0003	33	0.0026	50	0.0564	70	0.0014
26	0.0047	43	0.0544	63	0.0002	34	0.0038	51	0.0552	71	0.0010
27	0.0070	44	0.0495	64	0.0001	35	0.0054	52	0.0531	72	0.0007
28	0.0100	45	0.0440	65	0.0001	36	0.0075	53	0.0501	73	0.0005
29	0.0139	46	0.0382			37	0.0102	54	0.0464	74	0.0003
30	0.0185	47	0.0325			38	0.0134	55	0.0422	75	0.0002
31	0.0238	48	0.0271			39	0.0172	56	0.0377	76	0.0001
32	0.0298	49	0.0221			40	0.0215	57	0.0330	77	0.0001
33	0.0361	50	0.0177			41	0.0262	58	0.0285	78	0.0001
34	0.0425	51	0.0139			42	0.0312	59	0.0241		

附表 A.2 标准正态分布表

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

附表 A.3 χ^2分布表

$$P\{\chi^2 > \chi^2_\alpha(n)\} = \alpha$$

n \ α	0.995	0.99	0.975	0.95	0.90	0.75	0.25	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	0.102	1.323	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.575	2.773	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	1.213	4.108	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	1.923	5.385	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	2.675	6.626	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	3.455	7.841	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	4.255	9.037	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	5.071	10.219	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	5.899	11.389	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	6.737	12.549	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	7.584	13.701	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	8.438	14.845	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	9.299	15.984	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	10.165	17.117	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	11.037	18.245	22.307	24.966	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	11.912	19.369	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	12.792	20.489	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	13.675	21.605	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	14.562	22.718	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	15.452	23.828	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	16.344	24.935	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	17.240	26.039	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	18.137	27.141	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	19.037	28.241	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	19.939	29.339	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	20.843	30.435	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	21.749	31.528	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	22.657	32.620	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	23.567	33.711	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	24.478	34.800	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	25.390	35.887	41.422	44.985	48.232	52.191	55.003
32	15.134	16.362	18.291	20.072	22.271	26.304	36.973	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	27.219	38.058	43.745	47.400	50.725	54.776	57.648
34	16.501	17.789	19.806	21.664	23.952	28.136	39.141	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.465	24.797	29.054	40.223	46.059	49.802	53.203	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	29.973	41.304	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	30.893	42.383	48.363	52.192	55.668	59.892	62.883
38	19.289	20.691	22.878	24.884	27.343	31.815	43.462	49.513	53.384	56.896	61.162	64.181
39	19.996	21.426	23.654	25.695	28.196	32.737	44.539	50.660	54.572	58.120	62.428	65.476
40	20.707	22.164	24.433	26.509	29.051	33.660	45.616	51.805	55.758	59.342	63.691	66.766
41	21.421	22.906	25.215	27.326	29.907	34.585	46.692	52.949	56.942	60.561	64.950	68.053
42	22.138	23.650	25.999	28.144	30.765	35.510	47.766	54.090	58.124	61.777	66.206	69.336
43	22.859	24.398	26.785	28.965	31.625	36.436	48.840	55.230	59.304	62.990	67.459	70.616

附表 A.4 t 分布表

$$P\{T > t_\alpha(n)\} = \alpha$$

n \ α	0.25	0.10	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3138	12.7062	31.8207	63.6574
2	0.8165	1.8856	2.9200	4.3207	6.9646	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0322
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9028
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5177	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500

附表 A.5 F 分布表

$$P\{F > F_\alpha(m,n)\} = \alpha$$

$$\alpha = 0.005$$

n \ m	1	2	3	4	5	6	8	12	24	∞
1	16 211	20 000	21 615	22 500	23 056	23 437	23 925	24 426	24 940	25 465
2	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.5	199.5
3	55.55	49.80	47.47	46.19	45.39	44.84	44.13	43.39	42.62	41.83
4	31.33	26.28	24.26	23.15	22.46	21.97	21.35	20.70	20.03	19.32
5	22.78	18.31	16.53	15.56	14.94	14.51	13.96	13.38	12.78	12.14
6	18.63	14.45	12.92	12.03	11.46	11.07	10.57	10.03	9.47	8.88
7	16.24	12.40	10.88	10.05	9.52	9.16	8.68	8.18	7.65	7.08
8	14.69	11.04	9.60	8.81	8.30	7.95	7.50	7.01	6.50	5.95
9	13.61	10.11	8.72	7.96	7.47	7.13	6.69	6.23	5.73	5.19
10	12.83	9.43	8.08	7.34	6.87	6.54	6.12	5.66	5.17	4.64
11	12.23	8.91	7.60	6.88	6.42	6.10	5.68	5.24	4.76	4.23
12	11.75	8.51	7.23	6.52	6.07	5.76	5.35	4.91	4.43	3.90
13	11.37	8.19	6.93	6.23	5.79	5.48	5.08	4.64	4.17	3.65
14	11.06	7.92	6.68	6.00	5.56	5.26	4.86	4.43	3.96	3.44
15	10.80	7.70	6.48	5.80	5.37	5.07	4.67	4.25	3.79	3.26
16	10.58	7.51	6.30	5.64	5.21	4.91	4.52	4.10	3.64	3.11
17	10.38	7.35	6.16	5.50	5.07	4.78	4.39	3.97	3.51	2.98
18	10.22	7.21	6.03	5.37	4.96	4.66	4.28	3.86	3.40	2.87
19	10.07	7.09	5.92	5.27	4.85	4.56	4.18	3.76	3.31	2.78
20	9.94	6.99	5.82	5.17	4.76	4.47	4.09	3.68	3.22	2.69
21	9.83	6.89	5.73	5.09	4.68	4.39	4.01	3.60	3.15	2.61
22	9.73	6.81	5.65	5.02	4.61	4.32	3.94	3.54	3.08	2.55
23	9.63	6.73	5.58	4.95	4.54	4.26	3.88	3.47	3.02	2.48
24	9.55	6.66	5.52	4.89	4.49	4.20	3.83	3.42	2.97	2.43
25	9.48	6.60	5.46	4.84	4.43	4.15	3.78	3.37	2.92	2.38
26	9.41	6.54	5.41	4.79	4.38	4.10	3.73	3.33	2.87	2.33
27	9.34	6.49	5.36	4.74	4.34	4.06	3.69	3.28	2.83	2.29
28	9.28	6.44	5.32	4.70	4.30	4.02	3.65	3.25	2.79	2.25
29	9.23	6.40	5.28	4.66	4.26	3.98	3.61	3.21	2.76	2.21
30	9.18	6.35	5.24	4.62	4.23	3.95	3.58	3.18	2.73	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.35	2.95	2.50	1.93
60	8.49	5.79	4.73	4.14	3.76	3.49	3.13	2.74	2.29	1.69
120	8.18	5.54	4.50	3.92	3.55	3.28	2.93	2.54	2.09	1.43

$\alpha=0.01$ 续表

n \ m	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98.49	99.01	99.17	99.25	99.30	99.33	99.36	99.42	99.46	99.50
3	34.12	30.81	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	9.89	9.47	9.02
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91
11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.59	3.16
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.17	2.80	2.36
22	7.94	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00

α=0.025　　续表

n \ m	1	2	3	4	5	6	8	12	24	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	956.7	976.7	997.2	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.37	39.41	39.46	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.54	14.34	14.12	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	8.98	8.75	8.51	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.76	6.52	6.28	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.60	5.37	5.12	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.90	4.67	4.42	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.43	4.20	3.95	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.10	3.87	3.61	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.85	3.62	3.37	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.66	3.43	3.17	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.51	3.28	3.02	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.39	3.15	2.89	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.29	3.05	2.79	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.20	2.96	2.70	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.12	2.89	2.63	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.06	2.82	2.56	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.01	2.77	2.50	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	2.96	2.72	2.45	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	2.91	2.68	2.41	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.87	2.64	2.37	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.84	2.60	2.33	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.81	2.57	2.30	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.78	2.54	2.27	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.75	2.51	2.24	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.73	2.49	2.22	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.71	2.47	2.19	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.69	2.45	2.17	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.67	2.43	2.15	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.65	2.41	2.14	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.53	2.29	2.01	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.41	2.17	1.88	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.30	2.05	1.76	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.19	1.94	1.64	1.00

α=0.05 续表

n \ m	1	2	3	4	5	6	8	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9	249.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25
∞	3.84	2.99	2.60	2.37	2.21	2.09	1.94	1.75	1.52	1.00

$\alpha=0.10$　　续表

n \ m	1	2	3	4	5	6	8	12	24	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	59.44	60.71	62.00	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.37	9.41	9.45	9.49
3	5.54	5.46	5.36	5.32	5.31	5.28	5.25	5.22	5.18	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.95	3.90	3.83	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.34	3.27	3.19	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	2.98	2.90	2.82	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.75	2.67	2.58	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.59	2.50	2.40	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.47	2.38	2.28	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.38	2.28	2.18	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.30	2.21	2.10	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.24	2.15	2.04	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.20	2.10	1.98	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.15	2.05	1.94	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.12	2.02	1.90	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.09	1.99	1.87	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.06	1.96	1.84	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.04	1.93	1.81	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.02	1.91	1.79	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.00	1.89	1.77	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	1.98	1.87	1.75	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	1.97	1.86	1.73	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.95	1.84	1.72	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.94	1.83	1.70	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.93	1.82	1.69	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.92	1.81	1.68	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.91	1.80	1.67	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.90	1.79	1.66	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.89	1.78	1.65	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.88	1.77	1.64	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.83	1.71	1.57	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.77	1.66	1.51	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.72	1.60	1.45	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.17	1.67	1.55	1.38	1.00

附录 B　常用术语的汉英对照

F 分布　F- distribution

F 检验　F-test

t 分布　t- distribution

T 检验　T-test

Z 检验　Z-test

贝叶斯公式　Bayes' formulas

备择假设　alternative hypothesis

边缘分布函数　marginal distribution function

标准差　standard deviation

伯努利大数定律　Bernoulli law of large numbers

伯努利试验　Bernoulli trials

泊松分布　Poisson distribution

不相关　uncorrelated

大数定律　law of large numbers

单因素方差分析　single factor ANOVA 或 one-way ANOVA

第二类错误　type Ⅱ error

第一类错误　type Ⅰ error

棣莫弗-拉普拉斯定理　De Moivre-Laplace theorem

点估计　point estimation

独立同分布的中心极限定理　independence and distribution central limit theorem

对立事件　opposite event

二维离散型随机变量　two dimensional discrete random variable

二维连续型随机变量　two dimensional continuous random variable

二维随机变量　two dimensional random variable

二项分布　binomial distribution

方差　variance

方差分析　analysis of variance (ANOVA)

方差分析表　the ANOVA summary table

χ^2 分布　chi-square distribution

分布函数　distribution function

分布律　probability function

概率　probability

概率乘法公式　multiplication formulas of probability

概率密度函数　probability density function

个体　individual

古典概率　classical probability

回归　regression

回归方程　regression equation

回归分析　analysis of regression

回归函数　regression function

回归曲(直)线　regression curve (line)

回归系数　regression coefficient

基本事件　basic event

极大似然估计　maximum likelihood estimation

假设检验　hypothesis testing

χ^2 检验　chi-square test

检验统计量　test statistic

简单随机样本　simple random sample

矩　moment

矩估计　moment estimation

拒绝域　rejection region

均匀分布　uniform distribution

控制　control

离散型随机变量　discrete random variable

离散型随机变量的边缘分布律　the marginal distribution law of discrete random variable

连续型随机变量　continuous random variable

连续型随机变量的边缘概率密度　the marginal probability density of continuous random variable

联合分布函数　joint distribution function

联合分布律　joint distribution law

联合概率密度　joint probability density

两点分布　two-point distribution

临界值　critical value

频率　frequency

切比雪夫不等式　the Chebyshev inequality

切比雪夫大数定律 Chebyshev law of large numbers
区间估计 interval estimation
全概率公式 formulas of total probability
散点图 scatter diagram (scatter graph)
上 α 分位点 upper percentile of α
事件的独立性 independence of the event
数学期望 expectation
随机变量 random variable
随机变量的独立性 independence of random variable
随机实验 random experiment
随机事件 random event
条件分布函数 conditional distribution function
条件分布律 conditional distribution law
条件概率 condition probability
条件概率密度 conditional probability density
统计分布 statistical distribution
统计量 statistic
无偏估计 unbiased estimate
显著性检验 significance test
显著性水平 significance level
相关系数 correlation coefficient
协方差 covariance
辛钦大数定律 Khinchine law of large numbers
样本 k 阶矩 sample moment of order k
样本标准差 sample standard deviation
样本方差 sample variance
样本均值 sample mean
样本空间 sample space
一元线性回归 simple linear regression
一致估计量 uniform estimator
依概率收敛 convergence in probability
因素(因子) factor
有效估计量 efficiency estimator
预报 prediction
原假设 original hypothesis
正态分布 normal distribution
指数分布 exponential distribution
置信度(置信水平) degree of confidence (confidence level)
置信区间 confidence interval
置信上限 confidence upper limit
置信下限 confidence lower limit
中心极限定理 central limit theorem
总平方和 total variation
总体 population
组间平方和 variation between groups
组内平方和 variation within groups
最小二乘法 method of least squares

习题答案

习 题 1

1. (1) $S=\{HHH, HHT, HTH, HTT, THH, TTH, THT, TTT\}$

$A=\{HHH, HHT, HTH, HTT\}$

$B=\{HHT, HTH, HTT, THH, TTH, THT\}$

$C=\{HHH, HHT, HTH, HTT, THH, TTH, THT\}$;

(2) $S=\{(1,2,3),(1,2,4),(1,2,5),(1,2,6),(1,3,4),(1,3,5),(1,3,6),$
$(1,4,5),(1,4,6),(1,5,6),(2,3,4),(2,3,5),(2,3,6),(2,4,5),$
$(2,4,6),(2,5,6),(3,4,5),(3,4,6),(3,5,6),(4,5,6)\}$

$A=\{(2,3,4),(2,3,5),(2,3,6),(2,4,5),(2,4,6),(2,5,6)\}$

$B=\{(1,3,5)\}$

$C=\{(1,2,3),(1,2,4),(1,2,5),(1,2,6),(1,3,4),(1,3,5),(2,3,4)\}$;

(3) $S=\{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),$
$(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),$
$(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),$
$(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),$
$(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),$
$(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$

$A=\{(1,3),(1,5),(3,1),(5,1)\}$

$B=\{(1,2),(1,4),(1,6),(2,1),(2,3),(2,5),(3,2),(3,4),(3,6),$
$(4,1),(4,3),(4,5),(5,2),(5,4),(5,6),(6,1),(6,3),(6,5)\}$

2. (1) 该生是三年级男生，但不是数学建模竞赛队员；

(2) 只有三年级男生参加数学建模竞赛；

(3) 只有男生参加数学建模竞赛

3. (1) 0.3；　(2) 0.07；　(3) 0.14；　(4) 0.9；　(5) 0.1；　(6) 0.83

4. (1) $A\subseteq B$ 时，$P(AB)$最大，为 0.5；

(2) 当 $P(A\cup B)=1$ 时，$P(AB)$最小，为 0.3

5. 0.6　　6. 证略　　7. $\frac{1}{12}$

8. (1) $\frac{5}{9}$；　(2) $\frac{1680}{9^4}$；　(3) $\frac{10\times 8^3}{9^5}$；　(4) $1-\frac{8^5}{9^5}-\frac{C_5^1 8^4}{9^5}$

9. (1) $\frac{C_{48}^9 C_4^4}{C_{52}^{13}}$；　(2) $\frac{C_{13}^3 C_{13}^2 C_{13}^5 C_{13}^3}{C_{52}^{13}}$

10. $\frac{C_{10}^5 C_6^4 C_4^2}{C_{20}^{11}}$　　11. $\frac{3}{28}$

12. (1) $\frac{C_{200}^{30} C_{800}^{170}}{C_{1000}^{200}}$；　(2) $1-\frac{C_{200}^{1} C_{800}^{199}}{C_{1000}^{200}}-\frac{C_{800}^{200}}{C_{1000}^{200}}$

13. $\frac{1}{7}$　　14. $\frac{2}{3}$　　15. (1) $\frac{3}{5}$；　(2) 0.22　　16. (1) 0.988；　(2) $\frac{29}{35}$

17. 0.75　　18. 0.328　　19. 0.936

20. 证略　　21. $\frac{40}{43}$

22. (1) 0.896；　(2) $\frac{5}{896}$　　23. $\frac{3}{14}$　　24. $\frac{19}{32}$

25. (1) $\frac{n(N+1)+mN}{(m+n)(M+N+1)}$；

(2) $\frac{1}{(m+n)(m+n-1)(M+N+2)}[n(n-1)(N+2)+m(m-1)N+2nm(N+1)]$

26. 0.942

27. (1) $1-\frac{C_{10}^6 (C_2^1)^6}{C_{20}^6}$；　(2) $\frac{C_{10}^1 C_9^4 (C_2^1)^4}{C_{20}^6}$；　(3) $\frac{C_{10}^2 C_8^2 (C_2^1)^2}{C_{20}^6}$

28. $\frac{\pi+2}{2\pi}$　　29. $\frac{3}{7}$　　30. 证略　　31. (1) 0.1402；　(2) 1

32. 5 局 3 胜　　33. 0.862 904　　34. $\frac{5}{21}$

35. 0.125 449；0.281 478 2；0.593 072 8

36. (1) $\frac{163}{450}$；　(2) $\frac{109}{287}$

习　题　2

1.

X	0	1	2	3
p_k	$\frac{1}{3}$	$\frac{2}{3}\cdot\frac{1}{3}$	$\left(\frac{2}{3}\right)^2\cdot\frac{1}{3}$	$\left(\frac{2}{3}\right)^3$

2.

X	1	2	3
p_k	$\frac{6}{10}$	$\frac{3}{10}$	$\frac{1}{10}$

$P\{X>3\}=0$

3.

X	1	2	3	…	n	…
p_k	0.4	0.6×0.4	$0.6^2\times0.4$	…	$0.6^{n-1}\times0.4$	…

4. (1) $\frac{12}{23}$；　(2) $\frac{12}{17}$　　5. $X\sim B(5,1-0.9^{10})$

6. (1) 0.2048；　(2) 0.057 92；　(3) 0.993 28

7. $\frac{53}{512}$；1 道　　8. $\frac{4e^{-2}}{15}$

9. (1) $F(x)=\begin{cases}0, & x<0\\ 1-p, & 0\leqslant x<1\\ 1, & x\geqslant 1\end{cases}$; (2) $F(x)=\begin{cases}0, & x<0\\ \frac{1}{3}, & 0\leqslant x<1\\ \frac{5}{9}, & 1\leqslant x<2\\ \frac{19}{27}, & 2\leqslant x<3\\ 1, & x\geqslant 3\end{cases}$, $\frac{19}{27}$, $\frac{2}{3}$

10. $\frac{21}{25}$ 11.

X	−1	1	4
p_k	0.3	0.5	0.2

12. (1) $A=0, B=\frac{1}{2}, C=2$; (2) $f(x)=\begin{cases}x, & 0\leqslant x<1\\ 2-x, & 1\leqslant x<2\\ 0, & \text{其他}\end{cases}$; (3) $\frac{1}{2}$

13. (1) $A=0.5$;

(2) $F(x)=P\{X\leqslant x\}=\begin{cases}0, & x<-\frac{\pi}{2}\\ \frac{1}{2}(\sin x+1), & -\frac{\pi}{2}\leqslant x<\frac{\pi}{2}\\ 1, & x\geqslant\frac{\pi}{2}\end{cases}$

14. (1) $\frac{1}{4}$; (2) $\frac{4}{9}$ 15. 0.625 16. (1) e^{-1}; (2) e^{-1}

17. (1) 0.5328, 0.6977; (2) $c=3$; (3) d 至多是 0.436

18. (1) 70, 14.81; (2) 0.984

19.

Y	−1	1	7
p_k	0.1	0.5	0.4

20. (1) $a=\frac{1}{4}$ $b=1$; (2) $F(x)=\begin{cases}0, & x<0\\ \frac{1}{8}x^2, & 0\leqslant x<2\\ x-\frac{x^2}{8}-1, & 2\leqslant x<4\\ 1, & 4\leqslant x\end{cases}$

(3) $\frac{1}{8}$; (4) $f_Y(y)=\begin{cases}\frac{1-y}{16}, & -3<y<1\\ \frac{7+y}{16}, & -7\leqslant y\leqslant -3\\ 0, & \text{其他}\end{cases}$

21. (1) $f_Y(y)=\begin{cases}\frac{1}{6}, & -2<y<4\\ 0, & \text{其他}\end{cases}$; (2) $f_Y(y)=\begin{cases}\frac{1}{2y}, & e^{-1}<y<e\\ 0, & \text{其他}\end{cases}$

22. $f_Y(y)=\begin{cases}\sqrt{\dfrac{2}{\pi}}e^{-\frac{y^2}{2}}, & y>0\\ 0, & y\leqslant 0\end{cases}$

23. (C)　　24. (A)

25. (1)

X	0	±1	±2	±3
p_k	$\frac{29}{91}$	$\frac{21}{91}$	$\frac{9}{91}$	$\frac{1}{91}$

；　(2) $\frac{31}{91}$

26. (1) $A=\frac{1}{2}$；　(2) $F(x)=\begin{cases}\dfrac{1}{2}e^{x}, & x<0\\ 1-\dfrac{1}{2}e^{-x}, & x\geqslant 0\end{cases}$

27. $P\{Y=k\}=C_4^k\left(\frac{1}{2}\right)^k\left(1-\frac{1}{2}\right)^{4-k}$，　$k=0,1,2,3,4$

28. (C)　　29.

Y	-1	2	3
p_k	0.2	0.5	0.3

30. $F_Y(y)=\begin{cases}0, & y\leqslant 0\\ y, & 0<y<1\\ 1, & y\geqslant 1\end{cases}$；$f_Y(y)=\begin{cases}1, & 0<y<1\\ 0, & \text{其他}\end{cases}$，即 $Y\sim U(0,1)$

31. 后一方案优于前一方案

习　题　3

1. (1)

X \ Y	0	1
0	$\frac{a^2}{(a+b)^2}$	$\frac{ab}{(a+b)^2}$
1	$\frac{ab}{(a+b)^2}$	$\frac{b^2}{(a+b)^2}$

；

(2)

X \ Y	0	1
0	$\frac{a(a-1)}{(a+b)(a+b-1)}$	$\frac{ab}{(a+b)(a+b-1)}$
1	$\frac{ab}{(a+b)(a+b-1)}$	$\frac{b(b-1)}{(a+b)(a+b-1)}$

2.

X \ Y	0	1	2	3
0	0	0	0	1/8
1	0	0	3/8	0
2	0	3/8	0	0
3	1/8	0	0	0

3. (1) $A=\frac{1}{2},B=\frac{1}{\pi}$,或 $A=-\frac{1}{2},B=-\frac{1}{\pi}$; (2) $\frac{9}{32}$

4. (1) $k=12$; (2) $F(x,y)=\begin{cases}(1-e^{-3x})(1-e^{-4y}), & x>0,y>0\\ 0, & \text{其他}\end{cases}$;

(3) $(1-e^{-1})e^{-1}$

5. (1) $k=15$; (2) $\frac{5}{64}$

6. (1) $F_X(x)=\begin{cases}1-e^{-0.5x}, & x>0\\ 0, & x\leqslant 0\end{cases}$, $F_Y(y)=\begin{cases}1-e^{-0.5y}, & y>0\\ 0, & y\leqslant 0\end{cases}$;

(2) $f(x,y)=\begin{cases}0.25e^{-0.5(x+y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases}$, $f_X(x)=\begin{cases}0.5e^{-0.5x}, & x>0\\ 0, & x\leqslant 0\end{cases}$,

$f_Y(y)=\begin{cases}0.5e^{-0.5y}, & y>0\\ 0, & y\leqslant 0\end{cases}$; (3) $e^{-0.1}$

7.

X \ Y	0	1	2	$p_{i\cdot}$
−1	0	1/3	1/12	5/12
0	1/6	0	0	1/6
1	5/12	0	0	5/12
$p_{\cdot j}$	7/12	1/3	1/12	

8. (1)

X	1	2	3	4
$p_{i\cdot}$	1/4	1/4	1/4	1/4

Y	1	2	3
$p_{\cdot j}$	1/2	1/3	1/6

(2)

Y|X=4	1	2	3
p	1/3	1/3	1/3

X|Y=3	1	4
p	1/2	1/2

9. $f(x|y)=\begin{cases}\frac{1}{1-|y|}, & |y|<x<1\\ 0, & \text{其他}\end{cases}$ 10. $\frac{47}{64}$

11. (1)

X \ Y	0	1
−1	1/4	0
0	0	1/2
1	1/4	0

;

(2) 不相互独立

12. (1) $f_X(x)=\begin{cases}\frac{3}{2}x^2, & -1<x<1\\0, & 其他\end{cases}$, $f_Y(y)=\begin{cases}e^{-y}, & y>0\\0, & y\leqslant 0\end{cases}$;　(2) 相互独立

13. (1) $\frac{2}{\pi-2}$;

(2) $f_X(x)=\begin{cases}\frac{1}{\pi-2}x^2\sin x, & 0\leqslant x\leqslant\frac{\pi}{2}\\0, & 其他\end{cases}$;　$f_Y(y)=\begin{cases}\frac{1}{\pi-2}y\cos y, & 0\leqslant y\leqslant\frac{\pi}{2}\\0, & 其他\end{cases}$;

(3) 不相互独立

14. (1)

X \ Y	1	2	3
1	0.2	0.12	0.08
2	0.15	0.09	0.06
3	0.1	0.06	0.04
4	0.05	0.03	0.02

;　(2) 0.74

15. $\alpha=\frac{1}{18},\beta=\frac{2}{9},\gamma=\frac{1}{6}$

16. (1) $f(x,y)=\begin{cases}\frac{1}{2}e^{-\frac{y}{2}}, & 0<x<1,y>0\\0, & 其他\end{cases}$;　(2) 0.1445

17. (1)

Z	-2	-1	0	1	2
p	0.09	0.12	0.34	0.2	0.25

;

(2)

W	-1	0	1
p	0.12	0.68	0.2

;

(3)

M	-1	0	1
p	0.09	0.16	0.75

;

(4)

N	-1	0	1
p	0.51	0.24	0.25

18. (1) 不相互独立;　(2) $f_Z(z)=\begin{cases}\frac{1}{2}z^2e^{-z}, & z>0\\0, & z\leqslant 0\end{cases}$

19. (1) $F_{\max}(z)=\begin{cases}0, & z<0\\z^4, & 0\leqslant z<1\\1, & z\geqslant 1\end{cases}$, $f_{\max}(z)=\begin{cases}4z^3, & 0<z<1\\0, & 其他\end{cases}$;

(2) $\frac{15}{16}$；　(3) $f_Z(z)=\begin{cases} z^3 & 0\leqslant z<1 \\ 1-(z-1)^3, & 1\leqslant z<2 \\ 0, & \text{其他} \end{cases}$

20. $f_Z(z)=\begin{cases} ne^{-n(z-\theta)}, & z>\theta \\ 0, & z\leqslant\theta \end{cases}$

21. 略

22.

X \ Y	1	2	3	4
1	1/10	1/10	1/10	1/10
2	1/10	1/10	1/10	0
3	1/10	1/10	0	0
4	1/10	0	0	0

23. (1) $f_X(x)=\begin{cases} 2x, & 0<x<1 \\ 0, & \text{其他} \end{cases}$，$f_Y(y)=\begin{cases} 1-y/2, & 0<y<2 \\ 0, & \text{其他} \end{cases}$；

(2) $f_Z(z)=\begin{cases} 1-\frac{z}{2}, & 0<z<2 \\ 0, & \text{其他} \end{cases}$

24. (1) $f(x,y)=\begin{cases} \frac{1}{x}, & 0<y<x<1 \\ 0, & \text{其他} \end{cases}$；　(2) $f_Y(y)=\begin{cases} -\ln y, & 0<y<1 \\ 0, & \text{其他} \end{cases}$；

(3) $1-\ln 2$

25. $f_Z(z)=\begin{cases} 0, & z\leqslant 0 \\ \frac{1}{2}(1-e^{-z}), & 0<z\leqslant 2 \\ \frac{1}{2}(e^2-1)e^{-z}, & z>2 \end{cases}$　26. $f_Z(z)=\begin{cases} e^{-z}, & z>0 \\ 0, & z\leqslant 0 \end{cases}$

27. $a=4^{\frac{1}{3}}$　28. $f(z)=\begin{cases} 0, & z\leqslant 0 \\ \frac{1}{2}, & 0<z\leqslant 1 \\ \frac{1}{2z^2}, & z>1 \end{cases}$

29. 证略　30. 证略　31. 证略

习 题 4

1. (1) 2,2；　(2) 1,79；　(3) 2.　2. $1,-2,\frac{7}{3}$

3. $\frac{3}{10}$　4. 1500　5. $\frac{3}{5}\pi$

6. (1) 2,0；　(2) $-\frac{1}{15}$；　(3) 5　7. (1) $\frac{4}{5}$；　(2) $\frac{3}{5}$；　(3) $\frac{1}{2}$

8. $-0.2, 2.76, 27.6$　　9. $0, 0.5, 0.5$

10. (1) $a=\frac{1}{4}, b=1, c=-\frac{1}{4}$；　(2) $\frac{(e^2-1)^2}{4}, \frac{e^2(e^2-1)^2}{4}$　　11. 8.857 35

12. (1) 0,0,0；　(2) 不相关,不独立　　13. 略

14. $\frac{7}{6}, \frac{7}{6}, \frac{11}{36}, \frac{11}{36}, -\frac{1}{36}, -\frac{1}{11}$　　15. (A)　　16. (1) $\frac{1}{4}$；　(2) $-\frac{2}{3}$

17. 33.64　　18. $\frac{1}{p}, \frac{(1-p)}{p^2}$

19. 当 $y=3500$t 时,$E(Y)$最大,为 8.25×10^3 万元

20. $\frac{5}{3}$　　21. 5　　22. 21　　23. $4333\frac{1}{3}$

24. $\frac{1}{3\pi\sqrt{5}}e^{-\frac{8}{15}\left(\frac{x^2}{3}+\frac{xy}{4\sqrt{3}}+\frac{y^2}{4}\right)}$　　25. 证略

习　题　5

1. $\frac{1}{2}$　　2. $\frac{8}{9}$　　3. 0.0062　　4. 0.5762　　5. 4996　　6. 0.9525

7. (1) 0.6826；　(2) 66 564　　8. 12　　9. $\frac{1}{12}$　　10. (C)　　11. 不可靠

12. 190

习　题　6

1. (1) $N\left(\mu, \frac{\sigma^2}{n}\right), N(0,1), t(n-1), \chi^2(n), \chi^2(n-1)$；　(2) $t(n)$；

(3) $\chi^2(18), 18$；　(4) $\binom{n}{k}p^k(1-p)^{n-k}$；　(5) (39.51, 40.49)

2. (1) (C)；　(2) (C)　　3. (1) $\frac{49}{36}\sigma^2$；　(2) $3(\sigma^2+3\mu^2)$

4. (1) $\frac{6}{7\sqrt{2\pi}\sigma}e^{\frac{-18\left(x-\frac{7}{6}\mu\right)^2}{49\sigma^2}}$；　(2) $\frac{1}{2\pi\sigma^2}e^{-\frac{(x-\mu)^2+(y-\mu)^2}{2\sigma^2}}$；

(3) $\frac{1}{3}(X_1+X_2+X_3), \max\{X_1, X_2, X_3\}, X_1+X_2-2\mu$ 是统计量；$\frac{1}{\sigma^2}(X_1^2+X_2^2+X_3^2)$ 不是统计量

5. 18.6167, 10.9088　　6. (1) 0.9918；　(2) 165　　7. 0.99　　8. 0.10　　9. $n, \frac{2n}{m}$

10. 1147, 7578.8889　　11. (1) $\hat{p}=\frac{\overline{X}}{m}$；　(2) $\hat{p}=\frac{\overline{X}}{m}$　　12. (1) $\hat{\lambda}=\overline{X}$；　(2) $\hat{\lambda}=\overline{X}$

13. $\hat{\theta}=\frac{1}{\overline{X}}-1, \hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln(1-X_i)}$　　14. (1) 均为 $\overline{X}$；　(2) 证略

15. (1) 证略；　(2) $\hat{\theta}=\overline{X}-\frac{1}{2}$　　16. (1) 证略；　(2) $c=\frac{\sigma_2^2}{\sigma_2^2+\sigma_1^2}$

17. (1) (5.608,6.392); (2) (5.5584,6.4416)
18. (1) (14.754,15.066); (2) (0.0188,0.1512)
19. (432.3055,482.695) 20. 证略
21. (1) σ^2; (2) 2
22. (1) (C); (2) (D)
23. $\hat{\theta}=\min\{x_1,\cdots,x_n\}$
24. $2(n-1)\sigma^2$
25. $\hat{\theta}=\frac{3-\overline{X}}{4}$, $\hat{\theta}=\frac{7-\sqrt{13}}{12}$
26. (1) $\hat{\beta}=\frac{\overline{X}}{\overline{X}-1}$; (2) $\hat{\beta}=\frac{n}{\sum_{i=1}^{n}\ln X_i}$
27. (1) $\hat{\theta}=\overline{X}$; (2) $\hat{\theta}=\frac{2n}{\sum_{i=1}^{n}\frac{1}{X_i}}$
28. (1) $e^{\mu+\frac{1}{2}}$; (2) (−0.98,0.98); (3) $(e^{-0.48},e^{1.48})$
29. (−0.8986,0.0186)
30. (1) (0.00788,0.01597); (2) (0.1588,24.0076)

习 题 7

1. 接受假设 2. 可以认为锰的熔点显著高于1250℃
3. 可以认为平均成绩为70分
4. (1) 接受含量不超过17.5 mg; (2) 接受含量不超过17.5 mg
5. 强度均值没有改变 6. 方差变化了 7. 新工艺方差不再是0.112^2
8. 该日维尼纶纤度标准差不正常 9. 接受 H_1
10. (1) 标准差不为0.048 cm; (2) 不能认为 $\sigma^2\leqslant 0.048^2$ 11. 证略
12. 两种药的浓度无显著不同 13. (1) 不可以认为 $\sigma_1^2=\sigma_2^2$; (2) 可以认为 $\sigma_1^2\leqslant\sigma_2^2$

习 题 8

1. 有显著差异 2. 有显著差异 3. 无显著差异 4. 有显著影响
5. (1) $\hat{y}=-39.02+0.63x$; (2) $\hat{\sigma}^2=66.81$; (3) $\hat{y}_0=59.16$
6. (1) $\hat{y}=-2.735+0.483x$,或$\hat{y}=67.3+0.483(x-145)$;
 (2) $\hat{\sigma}^2=0.934$; (3) 显著
7. (1) $\hat{y}=193.95+1.80x$; (2) 显著; (3) (255.988,366.004)
8. (1) $\hat{y}=10.28+0.304x$,约为0.304; (2) 显著; (3) 78.68,(77.47,79.89)
9. (1) 图略; (2) $\hat{y}=\frac{x}{0.1312+0.0823x}$; (3) 显著